VERSATILE NUCLEIC ACID HAIRPIN MOTIF FOR PROGRAMMING BIOMOLECULAR SELF-ASSEMBLY PATHWAYS

Inventors: Peng Yin, Pasadena, CA (US); Niles A. Pierce, South Pasadena, CA (US)

Assignee: California Institute of Technology, Pasadena, CA (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 257 days.

Appl. No.: 12/152,893
Filed: May 16, 2008

Prior Publication Data

Related U.S. Application Data
Provisional application No. 60/930,457, filed on May 16, 2007.

Int. Cl. C07H 21/00 (2006.01) C40B 50/04 (2006.01) C12Q 1/68 (2006.01) C12N 15/63 (2006.01) B82Y 5/00 (2011.01) C12N 15/10 (2006.01) C40B 50/02 (2006.01)

U.S. Cl. CPC, C12N 15/63 (2013.01); B82Y 5/00 (2013.01); C12N 15/10 (2013.01); C12N 15/1031 (2013.01); C12Q 1/6844 (2013.01)

Field of Classification Search
CPC ... C12N 15/63; C12N 15/10; C12N 15/1031; B82Y 5/00; C12Q 1/6844
USPC ... 435/6.1; 536/23.1; 506/24, 25
See application file for complete search history.

References Cited
U.S. PATENT DOCUMENTS
4,714,680 A 12/1987 Civin
4,905,204 A 10/1990 Civin
5,270,163 A 12/1993 Gold et al.
5,312,728 A 5/1994 Lizardi et al.
5,435,413 A 6/1995 Hogan et al.
5,563,256 A 10/1996 Chakrabarty et al.
5,579,793 A 12/1996 Gajewski et al.
5,643,741 A 7/1997 Tsukamoto et al.
5,677,136 A 10/1997 Simmons et al.
5,716,827 A 2/1997 Tsukamoto et al.
5,750,397 A 5/1998 Tsukamoto et al.
5,928,913 A 7/1999 Efstathiou et al.
5,989,823 A 11/1999 Jayasena et al.
6,255,469 B1 7/2001 Seeman et al.
6,264,825 B1 7/2001 Blackburn et al.
6,361,944 B1 3/2002 Mirkin et al.
6,361,945 B1 3/2002 Becker et al.
6,506,559 B1 1/2003 Fire et al.
6,555,367 B1 4/2003 Spencer et al.
8,105,778 B2 1/2012 Dirks et al.
8,124,751 B2 2/2012 Pierce et al.
8,241,854 B2 8/2012 Yin et al.
8,318,921 B2 11/2012 Pierce et al.
8,497,364 B2 7/2013 Pierce et al.
8,577,438 B2 11/2014 Yin
8,662,241 B2 2/2015 Yin et al.
8,862,582 B2 2/2015 Dirks et al.
2001/014445 A1 8/2001 Umovitz

FOREIGN PATENT DOCUMENTS
EP 0273085 7/1988

OTHER PUBLICATIONS

Primary Examiner — Shannon Janssen
Attorney, Agent, or Firm — Knobe, Martens, Olson & Bear, LLP

ABSTRACT
The present invention relates generally to programming of biomolecular self-assembly pathways and related methods and constructs. A versatile nucleic acid hairpin motif for programming biomolecular self-assembly pathways for a wide variety of dynamic functions, reaction graphs for specifying pathways, and methods of using the hairpin motif are provided.

9 Claims, 56 Drawing Sheets
(54 of 56 Drawing Sheet(s) Filed in Color)
References Cited

U.S. PATENT DOCUMENTS

2006/0035375 A1 2/2006 Head et al.
2010/021901 A1 1/2010 Yin et al.
2010/021904 A1 1/2010 Pierce et al.
2012/022140 A1 1/2012 Yin et al.
2012/022243 A1 1/2012 Yin et al.
2012/022244 A1 1/2012 Yin et al.
2012/0190835 A1 7/2012 Pierce et al.
2012/0251583 A1 10/2012 Rothenmund
2015/0046165 A1 1/2015 Pierce et al.


References Cited


Friedrich et al., “A Cellular Screening Assay to Test the Ability of PKR to Induce Cell Death in Mammalian Cells, Molecular Therapy, vol. 12, No. 5, pp. 969-975, Nov. 2005.


REFERENCES CITED

OTHER PUBLICATIONS


Nakano et al., “Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: t(GCtGGtN), t(GCtGtNgt), t(tGtCtGGtN), and t(tGtCtGtNgt),” Biochemistry, vol. 41, pp. 14281-14292, American Chemical Society, 2002.


References Cited

OTHER PUBLICATIONS


Bolt et al., Differential Reactivities of the mono- and di-epoxide of 1,3-butanediol. Toxicology 113 (1996): 294-296.


References Cited

OTHER PUBLICATIONS


Schute-Merker et al., “no tail (nt) is the zebrafish homologue of the mouse T (Brachury) gene.” Development 120 (1994) 1009-1015.


References Cited

OTHER PUBLICATIONS
Communication pursuant to Article 94(3) EPC dated Nov. 7, 2012 from Application No. 08755764.1, filed May 16, 2008.


Patel et al., Cancer Biology & Therapy 14: 8, 693-696; Aug. 2013.


* cited by examiner
Figure 1b
Figure 3c
Figure 6a

Target dynamic function:
Catalytic formation of a 3-arm DNA junction

(1) Pathway specification

Figure 6b

(2.1) Complementarity relationships

Figure 6c

(2.2) Clamping/padding

Figure 6d

(2.3) Dimensioning

lal = lbl = lcl = lx1 = lyl = lzl = 6 nt

(3) Sequence design
Figure 7a
Figure 12b
Figure 16a

Figure 16b

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
<th>Step 5</th>
<th>Step 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A+I</td>
<td>(A)</td>
<td>(A)+B</td>
<td>(AB)</td>
<td>(AB)+C</td>
</tr>
<tr>
<td>A+B</td>
<td>(A)</td>
<td>(A)+B</td>
<td>(AB)</td>
<td>(AB)+C</td>
<td>(ABC)</td>
</tr>
<tr>
<td>(AB)+C</td>
<td>(AB)</td>
<td>(AB)+C</td>
<td>(ABC)</td>
<td>(ABC)+D</td>
<td>(D)</td>
</tr>
<tr>
<td>(CD)</td>
<td>(CD)+A</td>
<td>(CDA)</td>
<td>(CDA)+B</td>
<td>(CDA)+B</td>
<td></td>
</tr>
</tbody>
</table>
Nucleated growth of a binary tree (part II)
Figure 20
Figure 21
Figure 23
Figure 27c
Figure 27d
Figure 28a
<table>
<thead>
<tr>
<th>Step 1</th>
<th>Step 2</th>
<th>Step 3</th>
<th>Step 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 S1 S1 S1 W T1 S6 S6 S6 S6 S6 T1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2 S2 S2 S2 W S7 S7 S7 S7 S7 W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3 S3 S3 S3 S8 S8 S8 S8 S8 T2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4 S4 S4 S4 S9 S9 S9 S9 S9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS (T1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1+W+T2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 28b**
Catalytic formation of duplex A-B

Figure 29a
VERSATILE NUCLEIC ACID HAIRPIN MOTIF FOR PROGRAMMING BIOMOLECULAR SELF-ASSEMBLY PATHWAYS

REFERENCE TO RELATED APPLICATIONS

The present application claims priority to U.S. Provisional Application No. 60/930,457, filed May 16, 2007, which is incorporated herein by reference in its entirety.

BACKGROUND

1. Field

The present invention relates generally to programming of biomolecular self-assembly pathways and related methods and constructs.

2. Background

Molecular self-assembly, a fundamental process underlying the replication and regulation of biological systems, has emerged as an important engineering paradigm for nanotechnology. For example, molecular nanotechnology uses positionally-controlled mechanosynthesis guided by molecular systems. Molecular nanotechnology involves combining physical principles demonstrated by the molecular machinery of life, chemistry, and other nanotechnologies with the systems engineering principles found in modern macroscale factories.

In biological systems, self-assembling and disassembling complexes of proteins and nucleic acids bound to a variety of ligands perform intricate and diverse dynamic functions. Attempts to rationally encode structure and function into synthetic amino and nucleic acid sequences have largely focused on engineering molecules that self-assemble into prescribed target structures without explicit concern for transient system dynamics. See, Butterfoss, G. L. & Kuhlman, Annu. Rev. Biophys. Biom. 35, 49-65 (2006); Seeman, N. C., Nature 421, 427-431 (2003).

Current protocols for self-assembling synthetic DNA nanostructures often rely on annealing procedures to bring interacting DNA strands to equilibrium on the free energy landscape. Winfree et al., Nature 394, 539-544 (1998); Shih et al., Nature 427, 618-621 (2004); Rothemund, Nature 440, 297-302 (2006). Self-assembly in biology proceeds isothermally and assembly kinetics are often controlled by catalysts. To date, synthetic DNA catalysts have been used to control the kinetics of the formation of DNA duplex structures. Turberfield, A. J. et al., Phys. Rev. Lett. 90, 118102 (2003); Bois, J. S. et al., Nucleic Acids Res. 33, 4090-4095 (2005); Green, S. J. et al., Biophys. J. 91, 2366-2375 (2006); Seelig, G. et al., J. Am. Chem. Soc. 128, 12211-12220 (2006). However, until now, it has remained challenging to mimic nature's ability to encode dynamic function in the design space of biomolecules. Thus, there is a need for systems that are designed to autonomously perform dynamic functions.

SUMMARY

In some embodiments, the present teachings provide methods and products that biomolecular self-assembly pathways. In some embodiments, these methods and products can be in initiating and/or carrying out dynamic functions. In some embodiments, methods and compositions disclosed herein are beneficial for providing, for example without limitation, bio-markers and smart therapeutics that can detect a disease marker and then activate a therapeutic biological pathway.

In some embodiments, a hairpin monomer for performing a dynamic function is provided. The hairpin monomer comprises a first domain comprising a first toehold and a first propagation region, wherein the first toehold is exposed such that it is available to hybridize to a portion of a first nucleic acid sequence complementary to the first domain, and wherein the first toehold is located at an end of the monomer; and a second domain comprising a second toehold, wherein the second toehold is hybridized to a portion of the first propagation region.

In some embodiments the first nucleic acid sequence can be complementary to the first domain comprises a portion of an initiator molecule. In some embodiments the second toehold is configured to become available to hybridize to a second nucleic acid sequence of a second monomer if the first domain hybridizes to said initiator molecule.

In some embodiments the second domain further comprises a second propagation region, wherein said second propagation region is complementary to a portion of a second monomer. In some embodiments the said second toehold is configured to become available to initiate hybridization of said second propagation region to said portion of said second monomer if the first domain hybridizes to said first nucleic acid sequence complementary to said first domain. In some embodiments the second propagation region comprises a portion of a single stranded hairpin loop. The second propagation region can be configured not to initiate hybridization to said second monomer if the first domain hybridizes to the first nucleic acid sequence complementary to said first domain.

In some embodiments, the first domain is an input domain and the second domain is an output domain. In some embodiments, a portion of the first propagation region and the second toehold comprise a portion of a duplex stem.

In some embodiments, the hairpin monomer further comprises a third domain comprising a third toehold and a third propagation region, wherein the third toehold is hybridized to a portion of the first propagation region, and the third propagation region is single stranded.

In some embodiments the first toehold is single stranded. In some embodiments the first domain and second domain are concatenated in the monomer. In some embodiments the first domain and second domain are distinct and do not overlap.

In some embodiments, a method for initiating a dynamic function is provided. The method comprises: providing a first hairpin monomer comprising a first domain comprising a first toehold and a first propagation region, wherein the first toehold is exposed such that it is available to hybridize to a portion of a first nucleic acid sequence complementary to the first domain, and wherein the first toehold is located at an end of the monomer; and a second domain comprising a second toehold and a second propagation region, wherein the second toehold is hybridized to a portion of the first propagation region; and providing a second hairpin monomer. In some embodiments, the second hairpin monomer can comprise an input domain comprising a third toehold and a third propagation region, wherein the third toehold is exposed and complementary to the second toehold of the first hairpin monomer; and an output domain.

In some embodiments, the method further comprises providing a third hairpin monomer. In some embodiments, the method further comprises providing a fourth hairpin monomer. In some embodiments, the method further comprises providing a fifth hairpin monomer.

In some embodiments, the method further comprises providing an initiator comprising the first nucleic acid sequence complementary to the first domain of the first hairpin monomer. In some embodiments, the initiator comprises two
domains, wherein each domain is complementary to the first domain of the first hairpin monomer. In some embodiments, three or more molecules of the first hairpin monomer are present on a substrate. In some embodiments, the molecules of the first hairpin monomer are arranged linearly at regular intervals along a nicked DNA duplex.

In some embodiments, the dynamic function is selected from the group consisting of catalytic formation of a branched junction, autocatalytic duplex formation by a cross-catalytic circuit, nucleated dendritic growth, and autonomous locomotion.

In some embodiments, a self-assembly system for performing a dynamic function is provided. The self-assembly system for performing a dynamic function comprises: a first hairpin monomer comprising a first domain comprising a first toehold and a first propagation region, wherein the first toehold is exposed, and wherein the first toehold is located at an end of the monomer; and a second domain comprising a second toehold and a second propagation region, wherein the second toehold is hybridized to a portion of the first propagation region, and a second hairpin monomer.

In some embodiments, the output domain of the second hairpin monomer is complementary to the first domain of the first hairpin monomer. In some embodiments, the self-assembly system further comprises an initiator, wherein the initiator comprises an output domain comprising a third toehold complementary to the first toehold, and wherein the output domain is complementary to the first domain of the first hairpin monomer. In some embodiments, the initiator comprises a second output domain complementary to the first domain of the first hairpin monomer. In some embodiments, three or more molecules of the first hairpin monomer are present on a substrate. In some embodiments, the molecules of the first hairpin monomer are arranged linearly at regular intervals along a nicked DNA duplex.

In some embodiments, the second hairpin monomer comprises an input domain comprising a third toehold and a third propagation region, wherein the third toehold is exposed and complementary to the second toehold of the first hairpin monomer; and an output domain comprising a fourth toehold and a fourth propagation region, wherein the fourth toehold is hybridized to a portion of the third propagation region.

In some embodiments, the self-assembly system further comprises a third hairpin monomer, wherein said third hairpin monomer comprises a second input domain comprising a fifth toehold and a second output domain. Preferably, the fifth toehold is exposed and complementary to the fourth toehold of the output domain of the second hairpin monomer, and a portion of the second output domain is hybridized to a portion of the second input domain. In some embodiments, the second output domain can be complementary to the first domain of the first hairpin monomer. Other embodiments can include additional hairpin monomers.

In some embodiments, the second hairpin monomer further comprises a second output domain comprising a fifth toehold and a fifth propagation region, wherein the fifth toehold is hybridized to a portion of the third propagation region of the input domain of the second hairpin monomer, and the fifth propagation region is single stranded.

In some embodiments, the self-assembly system further comprises a third hairpin monomer, wherein the third hairpin monomer comprises: a third input domain comprising a sixth toehold and a sixth propagation region, wherein the sixth toehold is exposed and complementary to the fifth toehold of the second output domain of the second hairpin monomer.

In some embodiments, a method for programming a molecular pathway for carrying out dynamic function is provided. The method comprises: providing a reaction graph representing the molecular pathway for the dynamic function, wherein the reaction graph comprises: an initiator node representing an initiator molecule, wherein said initiator node comprises an initiator port; and at least one monomer node, wherein each monomer node in said set represents a hairpin monomer and comprises: an input port; and at least one output port, wherein each port of a node corresponds to a domain of the corresponding hairpin monomer or initiator molecule, and each domain comprises a toehold, wherein a port is in an accessible state if the toehold of the corresponding domain is exposed, and wherein a port is in an inaccessible state if the toehold of the corresponding domain is sequestered, and wherein the reaction graph indicates each reaction between each node; and translating the reaction graph to hairpin monomers. The reactions between the nodes can be assembly or disassembly reactions, or both.

In some embodiments, the method further comprises: designing nucleic acid primary sequences for the hairpin monomers. In some embodiments the dynamic function is selected from the group consisting of catalytic formation of a branched junction, autocatalytic duplex formation by a cross-catalytic circuit, nucleated dendritic growth, and autonomous locomotion. In some embodiments the reaction graph comprises two or more monomer nodes.

In some embodiments, reaction graph representing a molecular program for a dynamic function is provided. The reaction graph comprises: an initiator node representing an initiator molecule comprising an initiator port; and at least one monomer node, wherein each monomer node represents a hairpin monomer and comprises: an input port; and at least one output port, wherein each port of a node corresponds to a domain of the corresponding hairpin monomer or initiator molecule, and each domain comprises a toehold, wherein a port is in an accessible state if the toehold of the corresponding domain is exposed or in an inaccessible state if the toehold of the corresponding domain is sequestered, and wherein the reaction graph indicates each assembly and/or disassembly reaction between each node.

**BRIEF DESCRIPTION OF THE DRAWINGS**

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

**FIGS.** 1a-f generally depict various aspects of one embodiment for programming biomolecular self-assembly pathways. **a.** secondary structure of a hairpin monomer. Colored lines represent strand domains; short black lines represent base pairs; arrowheads indicate 3' ends. The small letters represent sequence segments. **b.** Secondary structure mechanism illustrating assembly and disassembly reactions during catalytic duplex formation. Letters marked with an asterisk (*) are complementary to the corresponding unmarked letter. **c.** Abstraction of the motif A as a node with three ports (color/shade use is consistent with a). **d.** A reaction graph representing a molecular program executed schematically in b and e.e. **e.** Execution of the reaction graph of d.f. **f.** Hierarchical design process.

**FIGS.** 2a-f generally depict catalytic self-assembly of three-arm and four-arm branched junctions. The small letters represent sequence segments. Letters marked with an asterisk (*) are complementary to the corresponding unmarked letter. **a.** Reaction graph for three-arm junctions. **b.** Secondary structure mechanism. **c.** Agarose gel electrophoresis demonstrating catalytic self-assembly for the three-arm system with
FIGS. 3a-c generally depict autocatalytic duplex formation by a cross-catalytic circuit with exponential kinetics. a. Reaction graph. Multiple assembly arrows entering the same input port depict parallel processes on separate copies of the nodal species. b. Secondary structure mechanism. c. System kinetistics examined by fluorescence quenching. Formation of A×B is monitored by the increase in fluorescence resulting from increased spatial separation between the fluorophores (green star in b) and the quencher (black dot in b) at either end of A. Raw data for two independent reactions are displayed for each initiator concentration (20-nM hairpins). Single traces are shown for the controls containing only A and B or only A. Inset: linear fit of the 100% completion time against the logarithm of the relative concentration of [[I]₀=0.05]]x. High-concentration end points ([[I]₀=0.1x]) are excluded based on theoretical analysis; low-concentration end points ([[I]₀=0.001x]) are excluded because of signal poisoning by leakage.

FIGS. 4a–e generally depict triggered assembly of quantized binary molecular trees. a. Reaction graph. Multiple assembly arrows entering the same input port depict parallel processes on separate copies of the nodal species. b. Secondary structure mechanism. c. Agarose gel electrophoresis demonstrating triggered self-assembly. Lanes 1–6: the dominant reaction band shifts with the addition of each generation of hairpins. Subdominant bands are presumed to represent imperfect dendrimers. Lane 7: minimal conversion to reaction products in the absence of initiator. Hairpins A1, A2, B2 at 62.5 nM; the concentration doubles for each subsequent generation of hairpins. Initiator I at 50 nM. d. Linear relationship between amplification signal (putative G5 reaction product) and initiator for three independent experiments (cross, diamond, circle). e. AFM imaging of dendrimers for G–3, 4, and 5.

FIGS. 5a,f generally depict stochastic movement of a bipedal walker. a. Reaction graph. Bonds between output ports on I and input ports on A represent initial conditions. Static structural elements are depicted by grey line segments. b. Secondary structure mechanism depicting processive locomotion. c–f, Fluorescence quenching experiments measuring the proximity of the quenchers (black dots) on the walker feet to the fluorophores (coloured stars) decorating the track. Fitted curves (solid) are used to determine the time at which the minimum fluorescence was achieved (dashed vertical line) for each fluorophore. c. Bipedal walker with track labeled by fluorophores DOE (green star)→TAMRA (red)→FAM (blue) as in b. d. Monopodal walkers on the same track (ORANGE star)→TAMRA (pale green)→FAM (pale blue). e. Comparison of track lengths for bipedal and monopodal walkers (eighteen traces per walker type: three fluorophores, six experiments). f. Bipedal walker with track labeled TAMRA (red star)→JOE (green)→FAM (blue).

FIGS. 6a–f generally depict the procedure for designing the catalytic e-arm junction system. a. The desired dynamic behavior: catalytic formation of a 3-arm DNA junction. b. Reaction graph for catalytic formation of a 3-arm DNA junction. c. Translation of the reaction graph to the secondary structure hairpin monomers. d. Addition of clamping and padding segments. e. Dimensioning. f. Sequence design. Green dot=A (adenine), blue dot=C (cytosine), black dot=G (guanine) and red dot=I (thymine).

FIGS. 7a–d generally depict execution of the reaction graphs for catalytic 3-arm and 4-arm junction systems. a. Execution of the reaction graph for catalytic 3-arm junction systems. b. Execution of the reaction graph for catalytic 4-arm junction systems. Figs. 8a–d depict catalytic formation of a 4-arm DNA junction. a. Reaction graph. Note: green output ports do not serve as initiators for any downstream reaction, and are omitted here for simplicity. b. Secondary structure schematic of the reaction. c. Agarose gel electrophoresis demonstrates the catalytic formation of the 4-arm junction. d. AFM images of two 4-arm junctions. Scale bar, 10 nm.

FIGS. 9a–b depict AFM measurements of the 3-arm (a) and 4-arm (b) junctions depicted in FIG. 2 and FIGS. 8a–d. The small images are screenshots of the measurement section files. The distance between the two arrows is listed above the image.

FIGS. 10a–b depict large-field-of-view AFM images of the 3-arm (a) and 4-arm (b) junction systems.

FIGS. 11a–b depict catalytic formation of a k-arm junction. a. Reaction graph. b. Reaction schematics.

FIGS. 12a–d depict catalytic formation of a 6-arm junction. a. Reaction graph. b and c. Step-by-step reaction schematic.

FIGS. 14 depicts the detailed reaction schematic for the autocatalytic system of FIG. 3. The length of each segment is 6 nt. Green star, fluorophore; black dot, quencher.

FIGS. 15a–c depict a step-by-step reaction schematic for the autocatalytic system of FIG. 3.

FIGS. 16a–b depict a reaction schematic and stepping gel for the autocatalytic system. a. Reaction schematic. b, Native polyacrylamide gel electrophoresis demonstrates the step-by-step reaction depicted in FIG. 36. The symbol ( ) indicates annealing; + indicates 15 minutes reaction at room temperature.

FIG. 17 depicts the execution of the reaction graph for the nucleated dendritic growth system.

FIG. 18 depicts a reaction schematic of the nucleated dendritic growth system (part I). Step-by-step reaction schematic of the nucleated dendritic growth system, as described in FIG. 4. The lengths of segments x, x*, and y are 2 nt; the lengths of the other segments are 7 nt. The figure continues in FIG. 19.

FIG. 19 depicts a reaction schematic of the nucleated dendritic growth system (part II). Step-by-step reaction schematic of the nucleated dendritic growth system, as described in FIG. 4. The figure continues from FIG. 9.

FIG. 20 depicts an agarose gel electrophoresis demonstrating quantitative amplification.

FIG. 21 depicts AFM measurements of the G3/G4 dendrimers.

FIG. 22 depicts AFM measurements of the G5 dendrimers.

FIG. 23 depicts large-field-of-view AFM image of the G5 dendrimer system.

FIGS. 24a–b depict execution of the reaction graph for the autonomous walker system of FIG. 5.

FIG. 25 depicts a secondary structure schematic of the walker system of FIG. 5. Stairs represent fluorophores; black dots represent quenchers. The lengths of segments a, b, c, and d are 7 nt; the lengths of segments x and y are 2 nt.

FIGS. 26a–b depict the step-by-step secondary structure schematic for the autonomous walker system of FIG. 5. Reaction arrows corresponding to the processive sub-population of walkers are shown in blue.

FIGS. 27a–d depict a detailed secondary structure schematic for step 1 of FIG. 26a. Reaction arrows corresponding to the processive sub-population of walkers are shown in purple.
FIGS. 28a-b depict assembly of the walker system. a. Assembly procedure. b. Native agarose gel electrophoresis demonstrating the expected assembly of the system.

FIGS. 29a-b depict a fuel system for the walker system. a. Reaction schematic. b. Agarose gel electrophoresis demonstrates catalytic formation of the DNA duplex.

FIG. 30 depicts DNA hairpin synthesis by ligation. The circled P indicates a phosphate group, which is used for ligation by, for example, T4 ligase.

FIG. 31 depicts a schematic of DNA sequences and secondary structures for the catalytic 3-arm junction systems of FIG. 2.

FIG. 32 depicts a schematic of DNA sequences and secondary structures for the catalytic 4-arm junction systems of FIG. 8.

FIG. 33 depicts a schematic of DNA sequences and secondary structures for the autocatalytic system of FIG. 3.

FIG. 34 depicts a schematic of DNA sequences and secondary structures for the nucleated dendritic growth system of FIG. 4.

FIG. 35 depicts a schematic of DNA sequences and secondary structures for the fuels for the walker system of FIGS. 28a-b and 29a-b.

FIGS. 36a-c generally depict systems with multiple inputs and outputs. a. Node A has one initially accessible input port which controls three initially inaccessible output ports. b. Node A has an initially accessible pink input port and an initially inaccessible orange input port; these two input ports together control an initially inaccessible output port. c. In general, a node can have m inputs and n outputs.

DETAILED DESCRIPTION

Various embodiments disclosed herein are generally directed towards programming biomolecular self-assembly pathways for dynamic functions as well as related methods and structures.

The difficulty of engineering molecular machines capable of nanoscale autonomous assembly or locomotion has attracted significant interest in recent years. Yin, P. et al., Angew. Chem. Int. Ed. 43, 4906-4911 (2004); Tian, Y. et al., Angew. Chem. Int. Ed. 44, 4355-4358 (2005); Bath, J. et al., Angew. Chem. Int. Ed. 44, 4358-4361 (2005); Pei, R. et al., J. Am. Chem. Soc. 128, 12693-12699 (2006); Venkataraman et al., Nat. Nanotechnol. 2, 490-494 (2007). Previous attempts to rationally encode structure and function into synthetic amino and nucleic acid sequences have largely been limited to engineering molecules that self-assemble into prescribed target structures without explicit concern for transient system dynamics. For example, previously, DNA dendrimer target structures have been synthesized via sequential ligation of structural subunits (Li, Y. et al. Nat. Mater. 3, 38-42 (2004)). However, the methods and compositions described herein make it possible to encode dynamic function in the design space of biomolecules.

A new approach to diverse molecular self-assembly pathways has been developed based on the rewiring of complementarity relationships between modular domains in a versatile hairpin motif. Monomer and polymer sequences can be encoded with the reaction pathways by which self-assembly occurs. In some embodiments this allows them to perform dynamic functions without human intervention. By programming complementarity relationships between domains within the hairpin motif, systems can be engineered to exhibit a wide variety of dynamic behaviors. The modular programmability of the hairpin motif can be used to facilitate the conversion of conceptual dynamical system designs into physical molecular implementations, enabling new approaches to fabrication, amplification, and transport (see, Yin et al., Nature 451(7176), 318-322; Supplementary Information pages 1-49 (2008), which is incorporated herein by reference in its entirety).

The versatile hairpin motif can be used to implement a variety of dynamic functions through self-assembly pathways. In some embodiments, starting from a conceptual dynamic function, a molecular implementation can be realized in three steps: (1) pathway specification via a “reaction graph”; (2) translation of the reaction graph into a secondary structure mechanism using monomers having the hairpin motif (“hairpin monomers”); and (3) computational design of hairpin monomer primary sequences.

Methods and compositions for programming biomolecular self-assembly pathways for dynamic functions, including, without limitation, molecular programming of catalytic geometry, catalytic circuitry, nucleated dendritic growth and autonomous locomotion are provided. In some embodiments, hairpin monomers to implement various dynamic functions are provided. In some embodiments, methods for designing hairpin monomers to implement dynamic functions are provided. In various embodiments, compositions and methods are provided for systems with catalytic geometry, catalytic circuitry, nucleated dendritic growth or autonomous locomotion.

One embodiment for programming a biomolecular self-assembly pathway is summarized in FIG. 1c. As shown in FIG. 1c, beginning with the dynamic function of, for example, the catalytic formation of a DNA duplex, a reaction graph using nodal abstractions is specified (step (1)). Next, the reaction graph is translated into motifs. In particular, hairpin monomers are designed encompassing the second structure mechanism of the functions provided by the reaction graph (step (2)). Then the primary sequences of the hairpin monomers are designed (step (3)).

A schematic depiction of the secondary structure of a hairpin monomer, which embodies the versatile hairpin motif, according to various embodiments disclosed herein is shown in FIG. 1b. The hairpin monomer (A) shown in FIG. 1a comprises three concatenated domains, a, b, and c. In other embodiments, monomers can comprise two or more concatenated domains. In preferred embodiments, each domain has a nucleation site called a toehold. For example, in FIG. 1a the toeholds are denoted a₀, b₀, and c₀ respectively. Preferably, a domain further comprises a propagation region. In some embodiments, the propagation region can be the portion of a domain that is not the toehold. Typically, a hairpin monomer comprises at least one input domain, and at least one output domain. In some embodiments, the input domain can be an initiator binding domain. In some examples, in FIG. 1b, domain a of hairpin monomer A is an initiator binding domain. In some embodiments, an output domain can be an assembly domain or a disassembly domain. For example, in FIG. 1b, domain b of hairpin monomer A is an assembly domain, and in domain a₀ of hairpin monomer B is a disassembly domain.

Two basic reactions can be programmed using the hairpin motif, as illustrated for one possible example of catalytic duplex formation in FIG. 1b. The reaction in FIG. 1b utilizes two hairpin monomers, A and B, each having two concatenated domains, a and b. First, an assembly reaction (1) occurs when a single-stranded initiator I containing an exposed toehold a⁺, nucleates at the exposed toehold a₀ of input domain a (also called the “initiator binding domain”) of hairpin monomer A, initiating a branch migration that opens the hairpin. Hairpin output domains b and c with newly exposed toeholds b₀ and c₀ can then serve as assembly initia-
tors for other suitably defined hairpins, permitting cascading (e.g., in reaction (2), output domain b (an "assembly domain") of hairpin monomer A assembles with input domain b* (an "assembly complement domain") of hairpin monomer B, opening the hairpin). Second, a disassembly reaction (3) occurs when a single-stranded output domain a^- of B (a "disassembly domain") initiates a branch migration that displaces the initiator I from A. In this example, I catalyzes the formation of duplex A • B via a prescribed reaction pathway.

To assist in programming more complex reaction pathways, a hairpin monomer can be abstracted as a node with input and output ports, with the state of the ports being indicated as either accessible or inaccessible. For example, the hairpin monomer of FIG. 1a can be abstracted as a node with three ports (FIG. 1c): a triangular input port and two circular output ports. The shade/color use for the nodal abstraction in FIG. 1c is consistent with FIG. 1a. The state of each port is either accessible (open triangle/circle) or inaccessible (solid triangle/circle), depending on whether the toehold of the corresponding hairpin domain is exposed or sequestered. Functional relationships between ports within a node are implicit in the definition of the nodal abstraction corresponding to a particular motif (e.g., for the node of FIG. 1c, the output ports flip to accessible states if the input port is flipped to an inaccessible state through an interaction with a complementary upstream output port).

Depicting assembly reactions by, for example, solid arrows and disassembly reactions by dashed arrows (each directed from an output port to a complementary input port of a different node), reaction pathways can be specified abstractly in the form of a reaction graph, representing a program to be executed by molecules such as, for example, nucleic acid molecules.

A reaction graph provides a simple representation of assembly (and disassembly) pathways that can be translated directly into molecular executables: nodes represent motifs, ports represent domains, states describe accessibility, arrows represent assembly and disassembly reactions between complementary ports. For example, the reactions depicted in the secondary structure mechanism of FIG. 1b are specified using a reaction graph in FIG. 1d. Conventions for the reaction graphs disclosed herein are provided below. The initial conditions for the program are described via the state (accessible or inaccessible) of each port in a reaction graph. FIG. 1e depicts the execution of this reaction graph through cascaded assembly and disassembly reactions. An assembly reaction is executed when ports connected by a solid arrow are simultaneously accessible.

The hairpin monomer functions as a modular programmable kinetic trap, and rewiring the connections between nodes in the reaction graph corresponds to rewiring the connections between kinetic traps in the underlying free energy landscape. In the physical systems, metastable hairpins are initially caught in engineered kinetic traps; the introduction of initiator molecules begins a chain reaction of kinetic escapes in which the hairpin species interact via programmed assembly and, optionally, disassembly steps to implement dynamic functions. Preferably, the time scale of metastability for kinetically trapped molecules is longer than the time scale relevant for the execution of the program.

As will be appreciated by one of skill in the art, the ability to design and implement biomolecular self-assembly (and disassembly) pathways can have great benefit, especially for engineering functional mechanical systems at the molecular scale. For example, the methods and compositions disclosed herein are beneficial for, inter alia, smart materials (e.g., artificial drugs and self-healing structures), nanosensors (e.g., photosensors), nanofacturing, nanorobots (e.g., replicating nanorobots and medical nanorobots), utility fog, and phased-array optics. In some embodiments, methods and compositions disclosed herein are beneficial for, for example, without limitation, smart therapeutics that can detect a disease marker (e.g., mutant mRNA known to cause cancer) and then activate a therapeutic biological pathway (e.g., kill the cancer cell leaving healthy cells untouched).

The above and additional embodiments are discussed in more detail below, after a brief discussion of the definitions some of the terms used in the specification.

DEFINITIONS

The section headings used herein are for organizational purposes only and are not to be construed as limiting the described subject matter in any way. All literature and similar materials cited in this application, including but not limited to, patents, patent applications, articles, books, treatises, and internet web pages are expressly incorporated by reference in their entirety for any purpose. When definitions of terms in incorporated references appear to differ from the definitions provided in the present teachings, the definition provided in the present teachings shall control. It will be appreciated that there is an implied "about" prior to the temperatures, concentrations, times, etc. discussed in the present teachings, such that slight and insubstantial deviations are within the scope of the present teachings herein. In this application, the use of the singular includes the plural unless specifically stated otherwise. Also, the use of "comprise", "comprises", "comprising", "contain", "contains", "containing", "include", "includes", and "including" are not intended to be limiting. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive.

Unless otherwise defined, scientific and technical terms used herein shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include plurals and plural terms shall include the singular. Generally, nomenclatures utilized in connection with, and techniques of, cell and tissue culture, molecular biology, and protein and oligo- or polynucleotide chemistry and hybridization described herein are those well known and commonly used in the art. Standard techniques are used, for example, for nucleic acid purification and preparation, chemical analysis, recombinant nucleic acid, and oligonucleotide synthesis.

Enzymatic reactions and purification techniques are performed according to manufacturer's specifications or as commonly accomplished in the art or as described herein. The techniques and procedures described herein are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the instant specification. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual (Third ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. 2000). The nomenclatures utilized in connection with, and the laboratory procedures and techniques of described herein are those well known and commonly used in the art.

As utilized in accordance with the embodiments provided herein, the following terms, unless otherwise indicated, shall be understood to have the following meanings:

A "self-assembly pathway" is a series of reactions autonomously executed by monomers in the formation of a polymer. The self-assembly pathway comprises assembly, or polymer-
ization, of monomers. In some embodiments, the self-assembly pathway can also comprise one or more disassembly reactions.

The term “nucleic acid” refers to natural nucleic acids, artificial nucleic acids, analogs thereof, or combinations thereof. Nucleic acids may also include analogs of DNA or RNA having modifications to either the bases or the backbone. For example, nucleic acid, as used herein, includes the use of peptide nucleic acids (PNA). The term “nucleic acids” also includes chimeric molecules.

The term “hairpin” as used herein refers to a structure formed by intramolecular base pairing in a single-stranded polynucleotide ending in an unpaired loop (the “hairpin loop”). In various embodiments, hairpins comprise a hairpin loop protected by stems. For example, a hairpin can comprise a first stem region, a hairpin loop region, and a second stem region. The first and second stem regions can hybridize to each other and together form a duplex region. Thus, a stem region of a hairpin monomer is a region that hybridizes to a complementary portion of the same monomer to form the duplex stem of a hairpin.

The term “hairpin loop” refers to a single stranded region that loops back on itself and is closed by a single base pair.

“Interior loop” and “internal loop,” are used interchangeably and refer to a loop closed by two base pairs. The closing base pairs are separate by single stranded regions of zero or more bases. A “bulge loop” is an interior loop where one of the separated single-stranded regions is zero bases in length and the other is greater than zero bases in length.

An “initiator” is a molecule that is able to initiate the polymerization of monomers. Preferred initiators comprise a nucleic acid region that is complementary to the initiator binding domain of a monomer.

“Monomers” as used herein refers to individual nucleic acid oligomers. Typically, at least two monomers are used in self-assembly pathways, although three, four, five, six or more monomers may be used. Typically each monomer comprises at least one domain that is complementary to at least a portion of one other monomer being used for the self-assembly pathway. Monomers are discussed in more detail below. In some embodiments, a monomer can have a hairpin motif. A monomer having a hairpin motif is referred to as a “hairpin monomer.” In other embodiments, a monomer can be an initiator.

The term “domain” refers to a portion of a monomer comprising a sequence. Preferably, a domain of a hairpin monomer comprises a toehold and a propagation region. An “input domain” of a monomer refers to a domain that is configured to receive a signal which initiates a physical and/or chemical change, such as, for example, a conformational change, of the monomer. Preferably, the signal is binding of a complementary sequence to the domain, typically beginning at the toehold. In some embodiments, an input domain can be an initiator binding domain, an assembly complement domain, or a disassembly complement domain. An “output domain” of a monomer refers to a domain that is configured to confer a signal. Preferably, the signal is binding of a complementary sequence to an input domain. In some embodiments, an output domain is configured to confer a signal to an input domain of another monomer. In some embodiments, an output domain can be, for example, an assembly domain, or a disassembly domain. In some embodiments, an output domain can be present in an initiator.

A first monomer in a self-assembly pathway preferably has an initiator binding or input domain (e.g., domain a of monomer A in FIG. 1a) that is complementary to a portion of an initiator. The initiator binding domain preferably has an exposed toehold. Binding of the initiator to the initiator binding domain initiates the self-assembly pathway. An initiator binding domain is an input domain.

A monomer preferably has at least one output domain (e.g., domain b of monomer A in FIG. 1a) that is complementary to an input domain of another monomer. An output domain on a hairpin monomer is preferably only available to interact with the input domain of the other monomer when a self-assembly pathway has been started by the initiator. For example, the assembly domain of a first monomer becomes available to hybridize to the assembly complement domain of a second monomer when the first monomer has already hybridized to at least a portion of an initiator, as discussed in more detail below.

The term “nucleate” as used herein means to begin a process of, for example, a physical and/or chemical change at a discrete point in a system. The term “nucleation” refers to the beginning of physical and/or chemical nucleic acid sequence points in a system. In some embodiments, nucleation of a self-assembly reaction can occur by, for example, the hybridization of a portion of an initiator to an exposed toehold of a hairpin monomer.

The term “toehold” refers to nucleation site of a domain comprising a nucleic acid sequence designed to initiate hybridization of the domain with a complementary nucleic acid sequence. The secondary structure of a monomer may be such that the toehold is exposed or sequestered. For example, in some embodiments, the secondary structure of the toehold is such that the toehold is available to hybridize to a complementary nucleic acid (the toehold is “exposed,” or “accessible”), and in other embodiments, the secondary structure of the toehold is such that the toehold is not available to hybridize to a complementary nucleic acid (the toehold is “sequestered,” or “inaccessible”). If the toehold is sequestered or otherwise unavailable, the toehold can be made available by some event such as, for example, the opening of the hairpin of which it is a part. When exposed, a toehold is configured such that a complementary nucleic acid sequence can nucleate at the toehold. In some embodiments, nucleation of a complementary nucleic acid sequence at an exposed toehold initiates branch migration that opens up the hairpin of a hairpin monomer.

A “propagation region” as used herein refers to a portion of a domain of a hairpin monomer that is configured to hybridize to a complementary nucleic acid sequence of another hairpin monomer once the toehold of the domain nucleates at an exposed toehold of the other hairpin monomer. The propagation region of a hairpin monomer is configured such that an available complementary nucleic acid sequence does not nucleate at the propagation region; rather, the propagation region hybridizes to a complementary nucleic acid sequence only after nucleation at the toehold of the same domain.

In some embodiments, monomers can be “metastable.” That is, in the absence of an initiator they are kinetically disfavored from associating with other monomers comprising complementary regions.

As used herein, the terms “polymerization” and “assembly” are used interchangeably and refer to the association of two or more monomers, or one or more monomers and an initiator, to form a polymer. The “polymer” may comprise covalent bonds, non-covalent bonds or both. For example, in some embodiments three species of monomers can hybridize sequentially to form a polymer comprising a three-arm branched junction.

As herein term “disassembly” refers to the disassociation of an initiator or at least one monomer from a polymer.
or another monomer. For example, polymers can disassemble from polymers, and monomers can disassemble from polymers.

As used herein “reaction graph” refers to a representation of assembly (and, optionally, disassembly) pathways that can be translated into molecular executables.

As used herein the terms “flip” and “switch” are used interchangeably and refer to a change from one state (e.g., accessible) to another state (e.g., inaccessible).

As used herein, an “aptamer” is an oligonucleotide that is able to specifically bind an analyte of interest other than by base pair hybridization. Aptamers typically comprise DNA or RNA or a mixture of DNA and RNA. Aptamers may be naturally occurring or made by synthetic or recombinant means. The aptamers are typically single stranded, but may also be double stranded or triple stranded. They may comprise naturally occurring nucleotides, nucleotides that have been modified in some way, such as by chemical modification, and unnatural bases, for example 2-aminopurine. See, for example, U.S. Pat. No. 5,840,867. The aptamers may be chemically modified, for example, by the addition of a label, such as a fluorophore, or a by the addition of a molecule that allows the aptamer to be crosslinked to a molecule to which it is bound. Aptamers are of the same “type” if they have the same sequence or are capable of specific binding to the same molecule. The length of the aptamer will vary, but is typically less than about 100 nucleotides.

System Design

Starting from a conceptual dynamic function, molecular implementation of a self-assembly pathway can be realized in three steps as summarized in, for example, FIG. 1f.

Step 1: Pathway specification. In some embodiments, the pathway that implements a target dynamic function can be specified using a reaction graph, discussed in detail below. A reaction graph provides a simple representation of assembly (and disassembly) pathways that can be translated into molecular executables. For example, nodes in the reaction graph represent hairpin monomers, ports of the nodes represent domains, states of the ports describe accessibility, and arrows between the nodes represent assembly and disassembly reactions between complementary ports. An assembly reaction is executed when ports connected by a solid arrow are simultaneously accessible.

Step 2: Translation to motifs. The reaction graph can be directly translated to hairpin monomer secondary structures. In other words, hairpin monomer secondary structures can be modeled and designed based on the nodes and functional relationships represented in the reaction graph. For example, a node in a reaction graph can be translated into a hairpin monomer. The ports on the node can be translated into the domains of the hairpin monomer. More particularly, an input port on a node can be translated into an input domain of a hairpin monomer, and an output port can be translated into an output domain. The functional relationships between the nodes in the reaction graph can be translated into the functionality of the domains of the hairpin monomer. Initial dimensioning of the number of nucleotides in each segment can be performed using, for example, the NUPACK server, which models the behavior of strand species in the context of a dilute solution. Dirks et al., SIMA Rev. 49, 65-88 (2007). Several examples of hairpin monomer secondary structure design based on a reaction graph are provided below.

Step 3: Sequence design. The composition of the monomers is not limited to any particular sequences or number of bases, and is designed based on the particular dynamic function. A number of criteria can be used to design the monomers to achieve the desired properties. These include, for example and without limitation, sequence symmetry minimization, the probability of adopting the initiator secondary structure at equilibrium, the average number of incorrect nucleotides at equilibrium relative to the target structure, and hybridization kinetics. In some embodiments, primary sequences of the hairpin monomers can be designed by considering a suite of structures that punctuate the intended reaction pathway. In some embodiments, structures that explicitly preclude undesired off-pathway interactions (e.g., structures specifying the absence of an interaction between two strands that should not pair) are considered in designing the hairpin monomers.

In some embodiments, the sequences can be optimized computationally to maximize affinity and specificity for a desired structure by minimizing the average number of incorrectly paired bases at equilibrium (Dirks et al., Nucleic Acids Research, 32:1392-1403, 2004.) In some embodiments, the optimization can be performed primarily by computer software. (R. M. Dirks and N. A. Pierce. Nucleic acid sequence design software, unpublished. 2007; J. N. Zadeh and N. A. Pierce. Multi-objective nucleic acid sequence design software, unpublished. 2007.) In some embodiments, further manual optimization based on the same design metric can be performed for a subset of crucial target structures. Monomers are described in detail below. Several examples of hairpin monomer primary sequence design based on secondary structures are provided below.

The thermodynamic behavior of the sequences can be further analyzed using, for example, the NUPACK server (www.nupack.org). (Dirks et al., SIAM Rev. 49, 65-88, 2007; Zadeh et al., NUPACK: a web-based tool for the analysis and design of nucleic acid systems. In preparation, 2007.) Stochastic kinetic simulation (Flamm et al., RNA, 6:325-338, 2000) can also be performed to confirm the absence of significant kinetic traps along the target reaction pathways. (J. M. Schäfer and E. Winfree. Multi-stranded kinetic simulation software, unpublished).

The physical self-assembly pathway system components (e.g., hairpin monomers) can be prepared using standard methods, including, for example, commercially available nucleic acid synthesizers or obtained from commercial sources such as Integrated DNA Technologies (Coralville, Iowa). The monomers and polymers can be verified using, for example and without limitation, gel electrophoresis, bulk fluorescence quenching, or single-molecule atomic force microscopy (AFM), discussed below.

Reaction Graph Conventions

This section provides conventions for the reaction graphs described and depicted herein. Of course, as will be appreciated by those of skill in the art, reaction graphs representing the self-assembly pathways described herein can be prepared using conventions other than those described below to achieve the same results.

Initial conditions. The initial condition of the system is defined by the state of each port and the initial bonds between the ports. An initial bond between an output port and an input port implies that an assembly reaction has already occurred prior to the execution of the reaction graph (see, e.g., the bond between the output port of I and the input port of A in FIG. 5a).

Static structural elements. Static structural elements are depicted by gray line segments (e.g., the substrate of FIG. 5a) and are inert during execution of the reaction graph. These elements can be used to impose geometric constraints on the execution of the reaction graph (e.g., the rigid substrate and inextensible torso of the walker system).

Execution starting points. Execution begins with any solid arrow (assembly reaction) connecting two accessible ports. In
a system lacking two accessible ports connected by a solid arrow, execution cannot begin (e.g., the removal of node I can prevent execution of the pathway).

Assembly reaction. An assembly reaction is depicted by a solid arrow that points from an input port to a complementary output port of a different node. An assembly reaction is executed when these two ports are simultaneously accessible. In the execution of an assembly reaction, a bond is formed between the two ports, they are flipped to their inaccessible states, and the internal logic of the node with the affected input port is applied to its output ports (e.g., for the present motif, the output ports are flipped to their accessible states). Multiple solid arrows entering the same input port depict parallel processes on separate copies of the nodal species (e.g., the input port of node A in FIG. 3a and the input ports of nodes A2-A5 and B2-B5 in FIG. 4a).

Disassembly reaction. A disassembly reaction is depicted by a dashed arrow that points from an input port to a complementary output port of a different node. Using nodal abstractions of the present hairpin motif, a disassembly arrow completes a disassembly cycle. For a cycle involving k nodes: input port 1 & blue output port 2 & input port 3 & blue output port 4 & blue output port 2k & input port 1, where & denotes reaction, & denotes a disassembly reaction, and & denotes the internal logical connection between the two ports on the same node. For example, FIG. 1d contains a disassembly cycle for k = 3; input port of A & blue output of A & input port of B & blue output of B & input port of A. FIG. 2a contains a disassembly cycle for k = 2; input port of A & blue output of A & input port of B & blue output of B & input port of A.

In physical terms, the displacing strand and the strand to be displaced emanate as adjacent branches for a k-arm junction, allowing nuclearization of the displacement branch migration (e.g., FIG. 2b). The special case of k = 2 corresponds to standard toehold-mediated strand displacement (e.g., FIG. 1b, where the whole of domain b of hairpin A serves as the toehold) (Yurke et al., Nature 406, 605-608 (2000)).

A disassembly reaction is executed when the participating input port is accessible and the participating input port is inaccessible (using nodal abstractions of the present motif; a disassembly arrow completes a disassembly cycle implies that the participating output port can only become accessible after the participating input port becomes inaccessible).

In the execution of a disassembly reaction (e.g., FIG. 1e), the existing bond from an (inaccessible) output port to an (inaccessible) input port is replaced by a new bond to the displacing (accessible) output port; the states of both output ports are flipped.

Multiple dashed arrows entering the same input port depict parallel disassembly cycles involving separate copies of the nodal species.

Reaction graphs can be extended beyond the present versatile motif by defining new nodal species that abstract the functional relationships between domains in other motifs. In some embodiments, the present hierarchical approach to rationally encoding dynamic function in nucleic acid sequences can be used in, for example, constructing a compiler for biomolecular function—an automated design process uses a modular conceptual system design as an input, and provides a set of biopolymer sequences that encode the desired dynamic system behavior as an output.

Nodal Abstractions

As discussed above, to assist in programming more complex reaction pathways, a hairpin monomer can be abstracted as a node with input and output ports, with the state of the ports being indicated as either accessible or inaccessible. For example, the hairpin monomer of FIG. 1a can be abstracted as a node with three ports (FIG. 1c): a triangular input port and two circular output ports. The color use for the nodal abstraction in FIG. 1c is consistent with FIG. 1a. The state of each port is either accessible (open triangle/circle) or inaccessible (solid triangle/circle), depending on whether the toehold of the corresponding hairpin domain is exposed or sequestered. Of course, the particular conventions used for the nodal abstractions can vary from those disclosed herein and achieve the same result. An initiator can also be abstracted as a node with an input port, with the state of the port being indicated as either accessible or inaccessible. In some embodiments, the node representing an initiator can be referred to as an initiator node. In some embodiments, the node representing an initiator can be referred to as an initiator node. In some embodiments, nodes can have multiple input ports and/or multiple output ports.

In the nodal abstractions, nodes represent hairpin monomers and initiators, ports represent domains, and the port states describe accessibility of the corresponding domains. For example, an input port represents an input domain, and an output port represents an output domain. In addition, an accessible port represents an exposed (accessible) domain, and an inaccessible port represents a sequestered (inaccessible) domain.

Functional relationships between ports within a node are implicit in the definition of the nodal abstraction corresponding to a particular motif (e.g., for the node of FIG. 1c, the output ports flip to accessible states if the input port is flipped to an inaccessible state through an interaction with a complementary upstream output port).

In some embodiments, nodal abstractions can be used in a reaction graph to model a dynamic function. Secondary structure mechanisms can then be modeled and designed based on the reaction graph. Hairpin monomer (and initiator) primary sequences can be designed from the secondary structure mechanisms.

Nodes having multiple input and output ports are shown in FIGS. 36a-c. In FIG. 36a, node A has one initially accessible input port which controls three initially inaccessible output ports. It implements the molecular logic: if A's input is rendered inaccessible by the arrival of node 1, then make all the three output ports accessible. The right panel depicts the molecular implementation using the hairpin motif. In FIG. 36a, toehold a1 is initially accessible, while toeholds b2 and c3 are inaccessible. Initiator I can hybridize with A and opens the hairpin, rendering toeholds b2 and c3 accessible. In FIG. 36b, node A has an initially accessible pink input port and an initially inaccessible orange input port; these two input ports control an initially inaccessible output port. It implements the molecular logic: if both of A's output ports are rendered inaccessible by the arrivals of nodes 11 and 12, then make the output port accessible. The right panel depicts the molecular implementation using the hairpin motif. In FIG. 36b, toeholds a1 and a2 are initially accessible; toehold b1 is initially inaccessible. Initiator I1 can hybridize with A, and renders a2 accessible; then the now accessible a2 can hybridize with b2, which opens the hairpin, rendering toehold b1 accessible. Generally, a node can have m inputs and n outputs (FIG. 36c). In preferred embodiments, at least one of the input ports is initially accessible; all the output ports are initially inaccessible.

The node implements a prescribed molecular logic such that only a prescribed combination or combinations of suitable activators can activate a corresponding combination of output ports. For example, if and only upon hybridizing with
US 9,217,151 B2

17

18

17 initiators I1 AND (I2 OR I3) but NOT (I4 OR I5), activates output ports (O1 AND O2 AND O4).

Reaction Graphs

Self-assembly reaction pathways can be specified abstractly in the form of a reaction graph, representing a program to be executed by molecules such as, for example, nucleic acid molecules. A reaction graph provides a simple representation of assembly (and disassembly) pathways that can be translated directly into molecular executables: nodes represent hairpin monomers, ports represent domains, states describe accessibility, arrows represent assembly and disassembly reactions between complementary ports. For example, the reactions depicted in the secondary structure mechanism of FIG. 1d are specified using a reaction graph in FIG. 1d. Conventions for the reaction graphs disclosed herein are provided above.

The initial conditions for the program are described via the state of each port in a reaction graph. For example, FIG. 1e depicts the execution of this reaction graph through cascaded assembly and disassembly reactions. An assembly reaction is executed when ports connected by a solid arrow are simultaneously accessible. For example, for the initial conditions depicted in FIG. 1d, the program starts with the execution of reaction (1). Reaction 1 (assembly): In an assembly reaction (executed here by the accessible output port of I and the complementary accessible input port of A), a bond is made between the ports and they are flipped, or switched, to accessible states; the two output ports of A are flipped to accessible states (based on the internal logic of node A). Reaction 2 (assembly): A bond is made between the newly accessible blue output port of A and the complementary accessible input port of B and both ports are flipped to inaccessible states; the output port of B is flipped to the accessible state (based on the internal logic of node B). Reaction 3 (disassembly): In a disassembly reaction (executed here by the newly accessible output port of B, the inaccessible input port of A, and the inaccessible output port of I), the bond between the output port of I and the input port of A is displaced by a bond between the output port of B and the input port of A; the states of the two output ports are flipped. FIG. 1f summarizes the hierarchical design process starting from a conceptual dynamic function, a molecular implementation can be realized in three steps: (1) pathway specification via a reaction graph; (2) translation into secondary structure hairpin monomers; (3) computational design of hairpin monomer primary sequences.

Examples of reactions graphs are provided below and include reaction graphs for various dynamic functions, including: catalytic geometry, catalytic circuitry, nucleated dendritic growth and autonomous locomotion.

Monomers

In some embodiments, the reaction graph can be used as a basis for modeling and designing a secondary structure mechanism using hairpin monomers. The reaction pathways by which self-assembly reactions occur are programmed within the primary sequences of the hairpin monomers.

A monomer having the hairpin motif (i.e., "hairpin monomers") typically has a hairpin structure having at least two distinct, concatenated domains. Typically, a hairpin monomer has at least one input domain and at least one output domain.

In preferred embodiments, each domain comprises a nucleation site called a toehold and a propagation region. Preferably, the toehold of a first input domain is exposed and thus available to hybridize to a complementary nucleic acid of another molecule. Preferably, the propagation region of the second domain comprises at least a portion of a hairpin loop region of the hairpin monomer. In preferred embodiments, the toehold of the second domain is hybridized to a portion of the propagation region of the first domain and therefore sequestered in the duplex stem of the hairpin and unavailable to hybridize to a complementary nucleic acid of another molecule. Displacement of the propagation region of the first domain from the toehold of the second domain exposes the toehold such that it is becomes available to hybridize with a complementary nucleic acid sequence of another nucleic acid, typically another monomer.

For example, in FIG. 1a, the monomer A comprises three concatenated domains, a, b and c. In FIG. 1b, domain a is an input domain, and domains b and c are output domains. In typical embodiments, a hairpin monomer comprises at least two distinct, concatenated domains. Typically, a hairpin monomer comprises at least one input domain. In some embodiments, a hairpin monomer can have one, two, three, four, five, six or more concatenated domains. Typically, a hairpin monomer has at least one output domain. In some embodiments, a hairpin monomer can have one, two, three, four, five, six or more input domains. In some embodiments in a hairpin monomer having more than one input domain, a first input domain toehold is exposed, and additional input domain toeholds are sequestered. In some embodiments, the sequestered input domain toeholds are sequestered by the duplex stem of the hairpin, and the corresponding domain propagation regions are located on bulge loops. Typically, a hairpin monomer can have one or two output domains. In some embodiments, a hairpin monomer can have one, two, three, four, five, six or more output domains. Typically, the toehold of an output domain is hybridized to a portion of a propagation region of an input domain and therefore sequestered. Displacement of the propagation region of the input domain from the toehold exposes the toehold such that it becomes available to hybridize with a complementary nucleic acid sequence of, generally, another monomer. In some embodiments, the hairpin monomer can have a second output domain. In some embodiments, the second output domain can comprise a single stranded region at an end of a hairpin monomer. For example, the second output domain can have a toehold which is hybridized to a portion of a propagation region of an input domain, and a single stranded propagation region.

Two or more distinct species of hairpin monomers are preferably utilized in a self-assembly pathway. Each monomer species typically comprises at least one domain that is complementary to a domain of another monomer species. However, the monomers are designed such that they are kinetically trapped and the system is unable to equilibrate in the absence of an initiator molecule that can disrupt the secondary structure of one of the monomers. Thus, the monomers are unable to polymerize in the absence of the initiator.

Introduction of an initiator species triggers a self-assembly pathway resulting in formation of one or polymers. In some embodiments the polymer comprises only a first and second monomer species. In other embodiments, the polymers can comprise additional nucleic acids. In the examples below, two or more hairpin monomers polymerize in the presence of an initiator to begin a self-assembly pathway. The self-assembly pathways disclosed herein are discussed in more detail below and include, for example, pathways for: catalytic geometry, catalytic circuitry, nucleated dendritic growth and autonomous locomotion. The self-assembly pathways typically result in formation of a polymer, such as, for example, a branched junction, an autocatalytic duplex, a binary molecular tree, or a bipedal walker.

A number of criteria can be used to design the monomers to achieve the desired properties. These include, for example and without limitation, sequence symmetry minimization, the probability of adopting the initiator secondary structure at
equilibrium, the average number of incorrect nucleotides at equilibrium relative to the target structure, hybridization kinetics, and the silencing target sequence. The composition of the monomers is not limited to any particular sequences or number of bases, and is designed based on the particular dynamic function. In some embodiments, the composition of the monomers can be designed based on the reaction graph and corresponding secondary structure mechanism of a particular dynamic function.

Monomers can be synthesized using standard methods, including commercially available nucleic acid synthesizers or obtained from commercial sources such as Integrated DNA Technologies (Coralville, Iowa). In some embodiments, the monomers can be DNA monomers. In other embodiments, the monomers can be RNA monomers. In some embodiments, the monomers can be RNA-DNA hybrids.

In some embodiments, at least two hairpin monomer species are utilized in a self-assembly pathway as illustrated in FIG. 1b. In the depicted embodiment, the monomers are denoted A and B. In FIG. 1b, monomer A comprises three concatenated domains: a, b, and c. Monomer B comprises two concatenated domains, represented by b* and c*. Each domain comprises a toehold and a propagation region. In FIGS. 1a and 1b, each domain comprises a single sequence segment; however, a domain can comprise any number of sequence segments. In some embodiments, a domain can comprise a portion of a sequence segment. In FIGS. 1a and 1b, the small letters with a subscript denote the toehold of the domain. For example in FIG. 1a, a denotes the toehold of domain a. In other embodiments, a toehold can comprise one or more sequence segments, or a portion of a sequence segment. For the example shown in FIG. 1b, the portion of a domain that is not the toehold is referred to as the “propagation region.” In FIGs. 1a and 1b, the small letters represent sequence segments, and letters marked with an asterisk (*) are complementary to the corresponding unmarked letter.

In preferred embodiments, the first stem region of a monomer can hybridize to the second stem region of the monomer to form the hairpin structure. In some embodiments, in the absence of an initiator, the first and second stem regions of each monomer are generally hybridized to form a duplex region of the monomer. The monomers each preferably comprise a hairpin loop region and two “stems regions”—a first stem region and a second stem region that are complementary and together can form a duplex region.

In the embodiment depicted in FIG. 1b, an initiator I comprises an output domain comprising a* having an exposed toehold a*. In the depicted embodiment, a first hairpin monomer A comprises an “initiator binding domain” (input domain a having an exposed toehold a), and a first “assembly domain” (output domain b having a sequestered toehold b). In the depicted embodiment, a second hairpin monomer B comprises a first “assembly complement domain” (input domain b* having an exposed toehold b*, and a “disassembly domain” (output domain a* having a sequestered toehold a*).

Assembly according to some embodiments of a self-assembly pathway having catalytic geometry is depicted in FIGS. 1b (1) and (2). A domain a* of the initiator I and the initiator binding domain a of the first hairpin monomer A are typically substantially complementary. That is, the domain a* of the initiator I is able to hybridize to the initiator binding domain a of the first hairpin monomer A, here a portion of domain a.

The initiator preferably comprises an exposed toehold. In FIG. 1b, the initiator I comprises an exposed toehold a*, which is a portion of the domain a*. Exposed toehold a* of the initiator is complementary to a domain a of a first hairpin monomer A. In some embodiments, the initiator binding domain of a first hairpin monomer can comprise an exposed toehold and at least a portion of the first stem region of the first hairpin monomer. For example, in the depicted figure, the first hairpin monomer A has an initiator binding domain a comprising the exposed toehold a, and a portion of the first stem region of A.

Preferably, upon hybridization of the initiator to the exposed toehold of the initiator binding domain of the first hairpin monomer, the second stem region is displaced from the first stem region. This opens the hairpin of the first hairpin monomer. For example, in FIG. 1b at (1), the initiator I nucleates at the exposed toehold a* of the hairpin monomer A by pairing a* with a. This induces a strand displacement interaction resulting in the hybridization of the initiator I at domain a* to the initiator binding domain a of the first hairpin monomer A to form the first complex (I*A).

In FIG. 1b, the first complex (I*A) has a newly exposed single-stranded tail that comprises the assembly domain b of the first hairpin monomer A. Monomer A also has another domain, domain c having toehold c, which is newly exposed. The assembly domain b has a newly exposed toehold b*.

In some embodiments, the assembly domain of a first hairpin monomer in the first complex can comprise a portion of the loop region and a portion of the second stem region of the first hairpin monomer. For example, in FIG. 1b, the assembly domain b of first hairpin monomer A comprises a portion of the loop region (the “propagation region” of b) and a portion of the second stem region of A (the toehold b*). In the absence of an initiator, the first and second stem regions of the first hairpin monomer are generally hybridized to form a duplex domain of the first hairpin monomer, and the first assembly domain of the first hairpin monomer is generally not available for hybridization to another monomer.

Preferably, upon hybridization of a newly-exposed toehold of the assembly domain of the first hairpin monomer to the exposed toehold of the assembly complement domain of the second hairpin monomer, the second stem region is displaced from the first stem region. This opens the hairpin of the second hairpin monomer. For example, in FIG. 1b, the exposed toehold b* of first hairpin monomer A in the first complex I*A nucleates at the exposed toehold b* of the second hairpin monomer B by pairing segment b* with b* (FIG. 1b at (2)). This induces a strand displacement interaction resulting in the hybridization of the first hairpin monomer A at the assembly domain b to the assembly complement domain b* of the second hairpin monomer B to form a second complex (I*A*B). In preferred embodiments, the exposed toehold of assembly complement domain of the second hairpin monomer is configured to nucleate at the newly exposed toehold of the assembly of the first monomer and not at the propagation region of the assembly domain of the first monomer. Preferably, the assembly complement domain of the second monomer the exposed toehold of assembly complement domain of the second monomer nucleates at the newly exposed toehold, thereby inducing a strand displacement interaction resulting in the hybridization of the assembly domain to the assembly complement domain.

In FIG. 1b, the second complex (I*A*B) has a newly exposed single-stranded tail that comprises the disassembly domain comprising the segment a* of the second hairpin monomer B.

In some embodiments, the disassembly domain of a second hairpin monomer can comprise a portion of the loop region and a portion of the second stem region of the third hairpin monomer. For example, in the depicted embodiment, the
disassembly domain a* of second hairpin monomer B comprises the loop region and a portion of the second stem region of B. In the absence of an exposed second assembly domain, the first and second stem regions of the second hairpin monomer are generally hybridized to form a duplex domain of the second hairpin monomer, and the disassembly domain of the second hairpin monomer is generally not available for hybridization to another monomer.

In other embodiments, instead of a disassembly domain, the second monomer can have a second assembly domain complementary to a second assembly complement domain of a third hairpin monomer. Any number of additional hairpin monomer species having one or more assembly domains can be used in a self-assembly pathway depending on the particular dynamic function.

In some embodiments, disassembly of an initiator from a monomer or polymer can occur. In some embodiments, polymers can disassemble from polymers, and monomers can disassemble from polymers. For example, disassembly of an initiator from a polymer can occur as generally depicted in FIG. 1b (3). The second hairpin monomer can have a disassembly domain which is substantially complementary to the initiator binding domain of the first hairpin monomer. In the depicted embodiment, the disassembly domain of the second substrate monomer B comprises a segment a* that is complementary to the initiator binding domain a of the first hairpin monomer of the first complex. The disassembly domain a* becomes accessible upon binding of the first assembly domain to the second hairpin monomer and opening of the hairpin of the second hairpin monomer (FIG. 1b (2)). Preferably, upon hybridization of a newly-exposed disassembly domain to the initiator binding domain of the first hairpin monomer, the initiator is displaced from the first hairpin monomer.

The system illustrated in FIGS. 1a-f and discussed above exhibits linear growth in response to initiator. However, in some embodiments, monomers can be designed to undergo triggered self-assembly into branched structures exhibiting quadratic growth or dendritic structures exhibiting exponential growth. See, Pierce et al., U.S. patent application Ser. No. 11/371,346, which is herein incorporated by reference in its entirety. In other embodiments, monomers can be designed to undergo autonomous locomotion of a bipedal walker, or other dynamic functions.

Exponential growth is limited by the available space such that it decreases to cubic amplification as the volume around the initiator fills. However, if chain reactions products are able to dissociate, exponential growth can be maintained until the supply of monomers is exhausted. In some embodiments, increasing the rate of polymer growth can enhance the ability to, for example, detect the presence of a low copy number target initiators, such as a single target molecule in a large test volume.

In some embodiments, the secondary structure is preferably such that the monomers are metastable under the reaction conditions in the absence of an initiator. In the presence of an initiator, the secondary structure of a first monomer changes such that it is able to hybridize to an exposed toehold of a second monomer species. This in turn leads to a change in the secondary structure of the second monomer, which is then able to continue the self-assembly pathway to form the desired structure.

Sequence segments of domains (for example, a, b, c, d, q, r, s, t, u, v, x, y and z as illustrated herein) are not limited to any particular sequences or number of bases, and are designed based on the particular dynamic function. In some embodiments, the primary sequence of the monomers can be designed based on the corresponding reaction graph and secondary structure mechanisms.

The length of the toeholds, propagation regions, hairpin loop regions, and stem regions of the monomers can be adjusted, for example to ensure kinetic stability in particular reaction conditions and to adjust the rate of polymerization in the presence of initiator. The hairpin loop regions are preferably between about 1 and about 100 nucleotides, more preferably between about 3 and about 30 nucleotides and even more preferably between about 4 and about 7 nucleotides. In some embodiments, the hairpin loop regions can be about 6 nucleotides in length and the stems are about 18 nucleotides long.

The toeholds can be located at any site on a hairpin monomer. The length of the toeholds can be adjusted, for example to ensure kinetic stability in particular reaction conditions and to adjust the rate of polymerization in the presence of initiator. The toeholds are preferably between about 1 and about 100 nucleotides, more preferably between about 3 and about 30 nucleotides and even more preferably between about 4 and about 7 nucleotides. In some embodiments, the toeholds of a hairpin monomer can be about 6 nucleotides in length and the stems are about 18 nucleotides long.

Several methods are available to reduce spurious monomer polymerization in the absence of initiator for dynamic functions including those with both higher order growth schemes and linear growth schemes. These include helix clamping, helix lengthening and loop entropy ratchets. In helix clamping, the single stranded regions in one or more of the monomers are truncated at each end so that the helices that they could potentially invade in other monomers are effectively clamped at the ends by bases that are not present in the single stranded regions. Experiments have shown that this can eliminate any spurious initiation. The amount of truncation that is effective to decrease or eliminate spurious initiation can be determined by routine experimentation. For example, control experiments can be performed using fluorescent gel electrophoresis time courses to monitor strand exchange between single stranded DNA and duplex DNA for different clamp lengths. Using spectrally distinct dyes for the initially single stranded DNA and for the two DNA species in the duplex allows independent monitoring of all species as strand exchange proceeds. These controls can provide a systematic basis for section of clamp dimensions.

In other embodiments utilizing hairpin monomers, loop entropy ratchets are used to reduce self-assembly in the absence of initiator. For example, in some embodiments, an initiator can open a hairpin monomer via a three-way branch migration. This reaction is reversible because the displaced strand is tethered in the proximity of the new helix. However, by increasing the length of the single-stranded loop, the entropy penalty associated with closing the loop increases. As a result, a longer loop will bias the reaction to proceed forward rather than returning to the uninitiated state. However, larger loops are more susceptible to strand invasion. To counter this effect and allow the use of larger loops, mismatches can be introduced between the loop sequences and the complementary regions of the other monomers. Again, the loop length and amount of mismatch that produces the desired reduction in non-specific self-assembly can be determined by the skilled artisan through routine experimentation.

Other refinements to the system stabilize the hairpin monomers to help prevent self-assembly in the absence of an initiator. This can be achieved, for example, via super-stable hairpin loop sequences (Nakano et al. Biochemistry 41:14281-14292 (2002), herein incorporated by reference in.
its entirety), with ostensible structural features that could further inhibit direct hybridization to the hairpin. In some embodiments, appending segments can be incorporated into the monomer to modulate the lengths of a hairpin monomer’s sticky-end, stem, and loop regions, permitting more flexible dimensioning. In other embodiments hairpin loops are made to be self-complementary at their ends. Thus, self-complementation “pinches” the hairpin loops, making them shorter. However, if the reactive exposed toeholds of each monomer are complementary to the hairpin loop regions on the opposite monomer, they will have a slight propensity to close up, thereby slowing down the reaction. This feature can be utilized if a slower reaction is desired. Completely self-complementary hairpins can also be used, for example if the monomer hairpins are forming dimers with interior loops that are more easily invaded than their hairpin counterparts.

In some embodiments, monomers can be derivatized with a compound or molecule, for example, to increase the molecular weight of the polymer resulting from execution of a self-assembly pathway. In some embodiments they can be derivatized at a location that does not interfere with their ability to hybridize. In some embodiments, the monomers comprise a protein-binding region, or other recognition molecule. In some embodiments, the monomers can contain a fluorophore, luminescent molecule, colorimetric compound or other component that allows the resulting polymers and/or the dynamic function to be visualized. Reaction conditions are preferably selected such that hybridization is able to occur, including between the initiator and the exposed toehold of a first hairpin monomer, between the assembly domain of a first hairpin monomer and an available toehold of a second hairpin monomer, and between the disassembly domain of a second hairpin monomer and the initiator binding domain, between the first and second stem regions of the monomers themselves. At each step of monomer polymerization, energy is gained from the hybridization of the exposed toehold of the monomer. The reaction temperature does not need to be changed to facilitate the polymerization of hairpin monomers. That is, hairpin monomer polymerization or assembly or disassembly reactions are isothermal. They also do not require the presence of any enzymes.

Initiators

As discussed above, an initiator can be a molecule that is able to initiate the polymerization of monomers. Typically, an initiator comprises an output domain that is complementary to an initiator binding domain (which is an input domain) of a hairpin monomer. In some embodiments, a self-assembly pathway is initiated when an initiator interacts with an initiator binding domain of a hairpin monomer, which subsequently undergoes a change in secondary structure, leading to polymer formation. In some embodiments, an initiator can initiate formation of a branched junction, an autocatalytic duplex, a binary molecular tree, or a bipedal walker.

Initiators can be synthesized using standard methods, including commercially available nucleic acid synthesizers or obtained from commercial sources such as Integrated DNA Technologies (Coralville, Iowa). Synthesis is discussed in more detail below. In some embodiments, initiators are naturally-occurring molecules. In other embodiments, initiators could be already present in a system. For example, the initiator may comprise nucleic acid naturally present in a system. In such as system the polymerization of monomers may be used to detect the presence of an initiator.

In some embodiments, the initiator binding domain of a first hairpin monomer is preferably at least 80%, more preferably at least 90%, 95% or higher, complementary to at least a portion of an initiator. In preferred embodiments, the initiator binding domain is at least 2, 3, 4, 5, or 10 or more bases in length.

The initiator preferably comprises a nucleic acid or other molecule that is able to contact a hairpin monomer and initiate a self-assembly pathway. In some embodiments, the initiator comprises a toehold having a sequence that is complementary to a portion, such as, for example without limitation, an exposed toehold, of a monomer, that is available for hybridization with the initiator while the monomer is in its kinetically stable state. In some embodiments, the initiator also preferably comprises a sequence that is complementary to a portion of the monomer adjacent to the toehold such that hybridization of the monomer to the toehold causes a conformational change in the monomer. For example, as depicted in FIG. 1b, the initiator I may comprise a toehold a° complementary to the initiator binding domain a of a monomer A, where the initiator binding domain a comprises a toehold a, and a portion of a first stem region of the monomer adjacent to the toehold.

In various embodiments, an initiator can be, for example without limitation, an RNA molecule, such as a coding region of RNA, a non-coding region of RNA, a portion of an mRNA, or a microRNA. In other embodiments, an initiator can be, for example without limitation, a DNA molecule such as, for example, a coding strand of DNA, or an antisense DNA.

In some embodiments, the initiator binding domain of a hairpin monomer can be a recognition molecule that specifically binds an initiator molecule. When the initiator interacts with the recognition molecule, the hairpin monomer undergoes a conformational change and the self-assembly pathway is initiated.

Recognition molecules include, without limitation, polypeptides, antibodies and antibody fragments, nucleic acids, aptamers, and small molecules.

In some embodiments, an initiator is bound to an aptamer, and the aptamer makes the initiator available to interact with a first hairpin monomer when the aptamer binds to an appropriate target molecule. For example, the initiator binding domain of a first hairpin monomer can bind to an initiator which is bound to an aptamer specific for a target of interest. Self-Assembly Pathways

Self-assembly pathways for a variety of different dynamic functions can be programmed via reaction graphs. The system illustrated in FIGS. 1a-1f and discussed above exhibits linear growth in response to initiator. However, a variety of different dynamic functions can be programmed using the methods disclosed herein. The programming of four different exemplary dynamic functions are generally described below:

1. catalytic formation of branched junctions,
2. autocatalytic duplex formation by a cross-catalytic circuit,
3. nucleated dendritic growth of a binary molecular tree, and
4. autonomous locomotion of a bipedal walker.

Catalytic Geometry

In some embodiments, the self-assembly pathway can be a pathway for catalytic formation of branched junctions. In some embodiments, a branched junction is formed in the presence of an initiator. The initiator may be, for example, any molecule in whose presence formation of a branch junction is desired. Initiators include, without limitation, polypeptides, such as antibodies and antibody fragments, nucleic acids, aptamers, and small molecules.

Compositions and methods are provided for catalyzing the formation of branched junctions. In some embodiments, the branched junction is a 3-arm, 4-arm or k-arm DNA junction (k≥3). For example, 3-arm DNA junctions are illustrated in Example 2, 4-arm DNA junctions are illustrated in Example
4, and k-arm junctions are illustrated in Example 6. The assembly and disassembly pathways for catalytic formation of a 3-arm DNA junction specified in the reaction graph of FIG. 2a are translated into the motif-based molecular implementation of FIG. 2b. The complementarity relationships between the segments of hairpins A, B, and C are specified (FIG. 2b, top) such that in the absence of initiator I, the hairpins are kinetically impeded from forming the three-arm junction that is predicted to dominate at equilibrium. In the reaction graph, this property is programmed by the absence of a starting point if node I is removed from the graph (i.e., no pair of accessible ports connected by an assembly arrow). The introduction of I into the system (FIG. 2b, bottom) activates a cascade of assembly steps with A, B, and C, followed by a disassembly step in which C displaces I from the complex, freeing I to catalyze the self-assembly of additional branched junctions. The design procedure for the catalytic 3-arm junction system shown in FIGS. 2a and b is described in detail below in the Examples section.

Each letter-labeled sequence segment shown in FIG. 2b is six nucleotides in length. However, as discussed above, sequence segments are generally not limited to any particular sequences or number of bases, and are designed based on the particular dynamic function. In FIG. 2b, the initially accessible toehold (a* for step (1)) or newly exposed toehold (b* for step (2); c* for step (3)) that mediates assembly reactions are labeled with purple letters.

In some embodiments, at least three hairpin monomers are utilized as illustrated in FIG. 2b to form a 3-arm junction. In FIG. 2b, the monomers are denoted A, B, and C. The monomers each preferably comprise an exposed toehold (for example, toeholds having sequences a, b, and c of A, B, and C, respectively), a hairpin loop region at the opposite end of the exposed toehold, and two “stems regions,” a first stem region and a second stem region, that together can form a duplex region. The small letters represent sequence segments. Letters marked with an asterisk (*) are complementary to the corresponding unmarked letter.

In preferred embodiments, the first stem region of a monomer can hybridize to the second stem region of the monomer to form the hairpin structure. For example, as shown in FIG. 2b, the monomer A comprises a first stem region comprising a sequence (x-y-b) that is able to hybridize to the second stem region (y*-b*-x*). In some embodiments, the absence of an initiator, the first and second stem regions of each monomer are generally hybridized to form a duplex region of the monomer.

In the embodiment depicted in FIG. 2b, an initiator I comprises a domain comprising the sequence a*-x*-b*-y*. In the depicted embodiment, a first hairpin monomer A comprises an “initiator binding domain” (comprising the sequence x-y) and a first “assembly domain” (comprising the sequence z*-c*-y*-b*). Typically, the initiator binding domain is an input domain and the assembly domain is an output domain. In the depicted embodiment, a second hairpin monomer B comprises a first “assembly complement domain” (comprising the sequence y-z-c-x) and a second “assembly domain” (comprising the sequence x*-a*-z*-c*). In the depicted embodiment, a third hairpin monomer C comprises a second “assembly complement domain” (comprising the sequence c-z-a-x) and a “disassembly domain” (comprising the sequence y*-b*-x*-a*). In the depicted embodiment, the assembly and disassembly domains are output domains, and assembly complement domains are input domains.

Assembly according to some embodiments of a self-assembly pathway having catalytic geometry is depicted in FIG. 2b (1)-(3). An output domain (a*-x*-b*-y*) of the initiator I and the initiator binding domain (a*-x*-b*-y*) of the first hairpin monomer A are typically substantially complementary. That is, the domain (a*-x*-b*-y*) of the initiator I is able to hybridize to the initiator binding domain (a*-x*-b*-y*) of the first hairpin monomer A.

The initiator I preferably comprises an exposed toehold a* which is a portion of the domain comprising the sequence a*-x*-b*. Exposed toehold a* of the initiator is complementary to a sequence segment a of a first hairpin monomer A. In some embodiments, the initiator binding domain of a first hairpin monomer can comprise an exposed toehold and a portion of the first stem region of the initiator. For example, in FIG. 2b, the first hairpin monomer A has an initiator binding domain a*-x*-b*-y, where a is an exposed toehold, and x*-b*-y is portion of the first stem region of the first hairpin monomer A.

Preferably, upon hybridization of the initiator to the exposed toehold of the initiator binding domain of the first hairpin monomer, the second stem region is displaced from the first stem region. This opens the hairpin of the first hairpin monomer. For example in FIG. 2b, the initiator I nucleates at the exposed toehold a* of the first hairpin monomer A by pairing segment a* with a (FIG. 2b at (1)). This induces a strand displacement interaction resulting in the hybridization of the initiator I at a domain a*-x*-b*-y* to the initiator binding domain a*-x*-b*-y* of the first hairpin monomer A to form the first complex (I*A).

In some embodiments, the first complex can have a newly exposed single-stranded tail that comprises a first assembly domain of the first hairpin monomer. For example, in FIG. 2b, the first complex (I*A) has a newly exposed single-stranded tail that comprises the first assembly domain (comprising the sequence z*-c*-y*-b*) of the first hairpin monomer A. This first assembly domain has a newly exposed toehold (b*).

In some embodiments, the first assembly domain of a first hairpin monomer in the first complex can comprise a portion of the loop region and a portion of the second stem region of the first hairpin monomer. For example, in FIG. 2b, the first assembly domain of first hairpin monomer A comprises the sequence z*-c*-y*-b*, where z*-c* is a portion of the loop region and y*-b* is a portion of the second stem region of the first hairpin monomer A. In the absence of an initiator, the first and second stem regions of the first hairpin monomer are generally hybridized to form a duplex domain of the first hairpin monomer, and the first assembly domain of the first hairpin monomer is generally not available for hybridization to another monomer.

Preferably, upon hybridization of a newly-exposed toehold of the first assembly domain of the first hairpin monomer to the exposed toehold of the first assembly complement domain of the second hairpin monomer, the second stem region is displaced from the first stem region. This opens the hairpin of the second hairpin monomer. In the depicted embodiment, the exposed toehold b* of first hairpin monomer A in the first complex I*A nucleates at the exposed toehold b of the second hairpin monomer B by pairing segment b* with b (FIG. 2b at (2)). This induces a strand displacement interaction resulting in the hybridization of the first hairpin monomer A at the first assembly domain z*-c*-y*-b* to the first assembly complement domain b*-y-c-z of the second hairpin monomer B to form a second complex (I*A+B).

In the depicted embodiment, the second complex (I*A+B) has a newly exposed single-stranded tail that comprises the second assembly domain (comprising the sequence x*-a*-z*-c*) of the second hairpin monomer B. This second assembly domain has a newly exposed toehold (c*).

In some embodiments, the second assembly domain of a second hairpin monomer can comprise a portion of the loop
region and a portion of the second stem region of the second hairpin monomer. For example, in FIG. 2b, the second assembly domain of second hairpin monomer B comprises the sequence γ*-b*-x*-a*-c*, where γ*-b* is a portion of the loop region and x*-a* is a portion of the second stem region of the second hairpin monomer. In the absence of an exposed first assembly domain, the first and second stem regions of the second hairpin monomer are generally hybridized to form a duplex domain of the second hairpin monomer, and the second assembly domain of the second hairpin monomer is generally not available for hybridization to another monomer.

Preferably, upon hybridization of a newly-exposed toehold of the second assembly domain of the second hairpin monomer to the exposed toehold of the second assembly complement domain of the third hairpin monomer, the second stem region is displaced from the first stem region. This opens the hairpin of the third hairpin monomer. For example, in FIG. 2b, the exposed toehold of the second hairpin monomer C in the second complex 1A*B nucleates at the exposed toehold c of the third hairpin monomer C by pairing segment c* with c (FIG. 2b at (3)). This induces a strand displacement interaction resulting in the hybridization of the second hairpin monomer B at the second assembly domain γ*-a*-z*-x*-c* to the second assembly complement domain c*-z*-a*-x of the third hairpin monomer C to form a third complex (1A*B*C).

In FIG. 2b, the third complex (1A*B*C) has a newly exposed single-stranded tail that comprises the disassembly domain (comprising the sequence γ*-b*-x*-a*) of the third hairpin monomer C. In some embodiments, the disassembly domain of the third hairpin monomer is complementary to a portion of the inhibitor binding domain of the first hairpin monomer.

In some embodiments, the disassembly domain of a third hairpin monomer can comprise a portion of the loop region and a portion of the second stem region of the third hairpin monomer. For example, in the depicted embodiment, the disassembly domain of third hairpin monomer C comprises the sequence γ*-b*-x*-a* where γ*-b* is a portion of the loop region and x*-a* is a portion of the second stem region of the third hairpin monomer. In the absence of an exposed second assembly domain, the first and second stem regions of the third hairpin monomer are generally hybridized to form a duplex domain of the third hairpin monomer, and the disassembly domain of the third hairpin monomer is generally not available for hybridization to another monomer.

In other embodiments for producing branched junctions with greater than 3 arms, instead of a disassembly domain, the third self-assembly has a third assembly domain complementary to a third assembly complement domain of a fourth hairpin monomer. For the formation of 3-arm branched junctions, preferably hairpin monomers are used. For the formation of 4-arm branched junctions, preferably four hairpin monomers are used. For the formation of k-arm branched junctions (where k≥3), preferably 2k hairpin monomers are used. In some embodiments, the kth hairpin monomer can comprise a disassembly domain instead of a kth assembly domain.

In some embodiments, the disassembly domain of the kth hairpin monomer is exposed by the opening of the hairpin of the kth hairpin monomer. The kth hairpin monomer has a disassembly domain which is substantially complementary to the initiator binding domain of the first hairpin monomer. The exposed disassembly domain can bind the initiator binding domain of the first hairpin monomer, thereby displacing the initiator from the kth complex such that the initiator can be recycled to react with another first hairpin monomer.

Disassembly according to some embodiments of a self-assembly pathway having catalytic geometry is depicted in FIG. 2b (4). The third hairpin monomer C has a disassembly domain which is substantially complementary to the initiator binding domain of the first hairpin monomer A. In the depicted embodiment, the disassembly domain of the third substrate monomer C comprises a sequence γ*-b*-x*-a* that is complementary to a sequence of the initiator binding domain a*-b*-y of the first hairpin monomer of the first complex (1A) that becomes accessible upon binding of the second assembly domain to the third hairpin monomer and opening of the hairpin of the third hairpin monomer (FIG. 2b (3)). Preferably, upon hybridization of a newly-exposed disassembly domain of the third hairpin monomer to the initiator binding domain of the first hairpin monomer, the initiator is displaced.

In some embodiments, the number of arms of a branch junction depends on the number of hairpin monomer species in the self-assembly pathway. The number and sequences of the hairpin monomers can be designed to provide branched junction having a desired sequence. In some embodiments, the result of the self-assembly pathway can be a 3-arm, 4-arm or k-branch branched junction.

In some embodiments, after displacement by the disassembly domain of the kth hairpin monomer, the displaced initiator can be used in further self-assembly reactions.

In some embodiments, gel electrophoresis can be used to confirm that the hairpins assemble slowly in the absence of initiator and that assembly is dramatically accelerated by the addition of initiator (FIG. 2c). Disassembly of the initiator enables catalytic turnover as indicated by the nearly complete consumption of hairpins even at substoichiometric initiator concentrations. Lanes 1-4 of the gel in FIG. 2c: show nearly complete conversion of hairpins to reaction products using stoichiometric or substoichiometric inhibitor I. In some embodiments, only minimal assembly is achieved by annealing the hairpin mixture, illustrating the utility of pathway programming for traversing free energy landscapes with kinetic traps that cannot be overcome by traditional annealing approaches. In FIG. 2c, minimal conversion is seen in the absence of initiator (lane 5), even with annealing (lane 6).

In some embodiments, direct imaging of the catalyzed self-assembly product (e.g., A*B*C) via atomic force microscopy (AFM) can be used to confirm the expected 3-arm junction morphology as shown in, for example, FIG. 2d.

In some embodiments, the reaction pathway can be extended to the catalytic self-assembly of k-arm junctions. FIGS. 2e and 2f show an example of catalytic self-assembly of a k-arm junction where k=4 in a reaction graph (FIG. 2e) and AFM image (FIG. 2f), respectively. A detailed example of the programming for the catalytic formation of a 4-arm junction is provided below in Example 4.

Catalytic Circuitry

In some embodiments, the self-assembly pathway can be a pathway for an autocatalytic system with exponential kinetics. Compositions and methods are provided for autocatalytic duplex formation by a cross-catalytic circuit. In some embodiments, the triggered exponential growth of cross-catalytic self-assembly pathway can be used in, for example, engineering enzyme-free isothermal detection methods. In sensing applications, self-replication can provide signal amplification for enzyme-free isothermal alternatives to polymerase chain reaction based on self-assembly reaction pathways.

In some embodiments, programming of the cross-catalytic self-assembly pathways can be executed as shown in FIGS. 3a and b. The reaction graph of FIG. 3a generates an auto-
catalytic system with exponential kinetics. In the corresponding molecular implementation (FIG. 3b), four hairpin species, A, B, C, and D coexist metastably in the absence of initiator I (FIG. 3b, top). The initiator catalyzes the assembly of hairpins A and B to form duplex A•B (Steps 1-2, FIG. 3b, bottom), bringing the system to an exponential amplification stage powered by a cross-catalytic circuit: the duplex A•B has a single-stranded region that catalyzes the assembly of C and D to form C•D (Steps 3-4); C•D in turn has a single-stranded region that is identical to 1 and can thus catalyze A and B to form A•B (Steps 5-6). Hence, A•B and C•D form an autocatalytic set capable of catalyzing its own production. Disassembly (Steps 2b, 4b, and 6b) is fundamental to the implementation of autocatalysis and sterically uninhibited exponential growth.

In some embodiments, each step in the reaction can be examined using, for example, native polyacrylamide gel electrophoresis to check for the expected assembly and disassembly behavior. System kinetics can be examined via a fluorescence quenching experiment (FIG. 3c). Spontaneous initiation in the absence of initiator reflects the finite time scale associated with the metastability of the hairpins and yields a sigmoidal time course characteristic of an autocatalytic system. As expected, the curve shifts to the left as the concentration of initiator is increased. A plot of 10% completion time against the logarithm of the concentration exhibits a linear regime, consistent with exponential kinetics and analytical modeling (FIG. 3c, inset). The minimal leakage of a system containing only A and B (labeled A•B in FIG. 3c) emphasizes that the sigmoidal kinetics of spontaneous initiation for the full system (A•B+C•D) are due to cross-catalysis.


A detailed example of one embodiment for programming catalytic circuitry is provided below in Example 7.

Nucleated Dendritic Growth

In some embodiments, the self-assembly pathway can be a pathway for nucleated dendritic growth. Compositions and methods for self-assembly pathways are provided in which nucleic acid monomers form dendrimers. In some embodiments, dendrimers are formed only upon detection of a target nucleation molecule. By growing to a prescribed size, such dendrimers can provide quantitative signal amplification with strength exponentially related to the number of constituent species.

In some embodiments, methods and compositions disclosed herein can be used for in situ amplification in bioimaging applications, such as, for example, in bio-marker generation. The bio-marker can, for example, facilitate fluorescence imaging, molecule sorting, etc. See, Pierce et al., U.S. patent application Ser. No. 11/371,346, which is herein incorporated by reference in its entirety. For example, monomers disclosed herein can be used to self-assemble a fluorescent polymer tethered to a target mRNA to detect expression of the mRNA. In some embodiments, labeled hairpin monomers can self-assemble in the presence of an initiator into a dendrimer of a prescribed size, yielding quantized signal amplification with strength exponential in the number of components. In some embodiments, only the root hairpin monomer and its two child species depend on the sequence of the initiator; thus, the other monomers do not need to be redesigned for each target initiator.

The molecular program in FIG. 4a depicts the triggered self-assembly of a binary molecular tree of a prescribed size. In the depicted embodiment, the reaction starts with the assembly of an initiator node 1 with a root node A1. Each assembled node subsequently assembles with two child nodes during the next generation of growth, requiring two new node species per generation. In the absence of steric effects, n-generation dendrimer uses 2(n-1) node species and yields a binary tree containing 2(n-1) monomers, i.e., a linear increase in the number of node species yields an exponential increase in the size of the dendrimer product. FIG. 4b depicts the motif based implementation of the program depicted in FIG. 4a: Hairpins are metastable in the absence of initiator; the initiator I triggers the growth of a dendrimer with five generations of branching (G5).

In one embodiment, trees with G=1, 2, 3, 4, and 5 are constructed. The nucleated growth of the trees can be examined using, for example, native agarose gel electrophoresis. Band shifting demonstrates increasing dendrimer size with each generation of growth (FIG. 4c). FIG. 4d demonstrates that the concentration of dendrimer depends linearly on the concentration of the initiator in the system. Finally, AFM imaging of dendrimers for G=3, 4, and 5 reveals the expected morphologies (FIG. 4e). Measurements of the dendrimer segment lengths agree well with the design.

A detailed example of one embodiment for programming nucleated dendritic growth is provided below in Example 9.

Autonomous Locomotion

In some embodiments, the self-assembly pathway can be a pathway for autonomous locomotion. Compositions and methods are provided for an autonomous enzyme-free bipolar DNA walker capable of stochastic locomotion along a DNA track. In some embodiments, the bipolar DNA walker system can mimic the bipolar motor protein, kinesin, which hauls intracellular cargo by striding along microtubules. Asbury et al., Curr. Opin. Cell Biol. 17, 89-97 (2005).

In some embodiments of a system for autonomous locomotion, they system comprises two species of “fuel” hairpin monomers, and a bipolar walker. In various embodiments, one species of fuel hairpin monomer (“track monomers”) can be linearly arranged at regular intervals along substrate to form a track for a bipolar walker. In some embodiments, the substrate can comprise a nicked DNA duplex. Typically, the track monomers comprise an input domain and an output domain. A bipolar walker comprises two identical “walker legs” connected by a duplex torso (FIG. 5a). The walker legs comprise an output domain complementary to an input domain of the track monomers. The bipolar walker assembles with the track monomers. In the presence of the second fuel hairpin monomer assemblies with a track monomer, which subsequently displaces from the bipolar walker. Typically, a second “fuel” hairpin monomer species comprises an input domain complementary to the output domain of the track monomer species, and an output domain complementary to the input domain of the track monomer species. The bipolar walker moves unidirectionally along the linear track by sequentially catalyzing the formation of a “track monomer-second fuel hairpin monomer” complex.

The molecular program in FIG. 5a depicts a self-assembly pathway the stochastic movement of a bipolar walker. Joined by a duplex torso, each of two identical walker legs, I, is capable of catalyzing the formation of a waste duplex A-B from metastable fuel hairpins A and B via a reaction pathway in which I assembles with A, which assembles with B, which
subsequently disassembles I from the complex. FIGS. 5a and b depict a reaction graph and corresponding molecular implementation for an exemplary bipedal walker. As shown in FIG. 5c, in some embodiments, the track can consist of five A hairpins arranged linearly at regular intervals along a nicked DNA duplex. In the presence of hairpin B, a sub-population of walkers will move unidirectionally along the track by sequentially catalyzing the formation of A•B. Due to the one-dimensional arrangement of anchor sites, this processive motion occurs only for those walkers that exhibit a foot-over-foot gait by stochastically lifting the back foot at each step.

One embodiment of a fuel system for a walker system is shown in FIG. 30a. Hairpins A and B in co-exist metastably in the absence of catalyst I. Catalyst I catalyzes A and B to form duplex A•B. Step 1: the toehold a of I nucleates at the toehold a of A, resulting in the opening of the hairpin A and the formation of the product I•A. Step 2: I•A, with e newly exposed, opens hairpin B; B subsequently displaces I from A, producing waste product A•B.

In some embodiments, walker locomotion can be investigated using a bulk fluorescence assay that tests whether there is a sub-population of walkers that locomotes processively in positions 3, 4, and 5, starting from an initial condition with legs anchored at positions 1 and 2. Quenchers are attached to the walker’s legs and spectrally distinct fluorophores are positioned proximal to anchorages 3, 4, and 5. Consistent with processivity, the anticipated sequential transient quenching of the fluorophores at positions 3, 4, and 5 is observed (FIG. 5c).

To rule out the possibility that this signal arises from non-processive walker diffusion through the bulk solution from one position to the next, monopodal walkers that lack a mechanism for achieving processivity can be used. In this case, the sequential transient quenching will no longer match the ordering of the fluorophores along the track (FIG. 5d) and the time scale for visiting any one of the three anchorages is longer than the time scale to visit all three anchorages for the bipedal system (FIG. 5c).

Additional control experiments show that this difference in time scales cannot be explained by the relative rates with which freely diffusing bipedal and monopodal walkers land on the track. As a further test of processivity for the bipedal walker, reordering the fluorophores along the track leads to the expected change in the ordering of the transient quenching (FIG. 5f).

A detailed example of one embodiment for programming autonomous locomotion is provided below in Example 11.

Pathway Analysis

The hairpin monomers, polymers, self-assembly pathway reactions and dynamic functions can be analyzed by any of a variety of methods known in the art. For example, gel electrophoresis can be used to compare the hairpin monomers before and after the reaction. For example, an amount of each monomer species can be mixed and mixed with varying amounts of initiator and a control (e.g., reaction buffer only). The samples can be allowed to react for a suitable time, such as for example without limitation 2 hours. The annealed can be mixed with loading buffer mix was loaded into a gel. The gel can be run and the nucleic acid visualized under UV light.

In some embodiments, the gel can be imaged using an imaging system such as, for example, an FLA-5100 imaging system (Fuji Photo Film Co., Ltd.).

In some embodiments, the hairpin monomers and polymers can be visualized using an atomic force microscope (AFM). For example, images can be obtained using a multimode scanning probe microscope, equipped with a Q-control module for analog AFMs. In some embodiments, samples can be first diluted in an appropriate buffer to achieve the desired sample density. The diluted sample can be applied onto the surface of freshly cleaved mica and allowed to bind.

Supplemental Ni++ can be added to increase the strength of DNA-nica binding. H. G. Hansma and D. E. Lane, *Biophysical Journal*, 70:1933-1938, 1996. Before placing the fluid cell on top of the mica puck, an additional amount of buffer can be added to the cavity between the fluid cell and the AMF cantilever chip to avoid bubbles.

In some embodiments, fluorescence data can be obtained for a walker system using, for example, a spectrofluorometer. For example, excitation and emission wavelengths were set to 394 nm and 517 nm (for FAM), 527 nm and 551 nm (for JOE), and 558 nm and 578 nm (for TAMRA), respectively, with 4 nm bandwidth. The assembly of the walker system is described above. In the experiments, an amount of the track and an amount of the bipedal walker can be used to assemble the system. In some embodiments, a substoichiometric amount of walker can be used to ensure that no free-floating walker would bind to the hairpin monomer on the track. For the same reason, a sub-stoichiometric amount of monopodal walker can be used in the diffusion experiments. The assembled track can be introduced first to record the three fluorescence baselines of FAM, JOE, and TAMRA. The hairpin monomer is then introduced to start the walker’s locomotion.

Compositions for Self-Assembly Pathways

Compositions and kits for self-assembly pathways are contemplated for use within the scope of the subject matter. In preferred embodiments, the compositions comprise a first hairpin monomer and a second hairpin monomer. In some embodiments, the compositions comprise a first hairpin monomer, a second hairpin monomer and a third hairpin monomer. In some embodiments, the compositions comprise a first hairpin monomer, a second hairpin monomer, a third hairpin monomer and a fourth hairpin monomer. Additional monomers can be included in some embodiments. In the presence of initiator, a self-assembly pathway is initiated causing the initiation of the desired dynamic function. In some embodiment, the dynamic function results in formation of a polymer. In some embodiments, in the presence of a catalyst, autonomous locomotion is initiated.

The compositions can also contain other components, such as, for example, accessory molecules that facilitate initiator recognition and aid the formation of polymers. Accessory molecules typically comprise nucleic acid molecules. In some embodiments, the accessory molecules are DNA helper strands that bind to regions flanking an initiator nucleic acid sequence. Preferably the accessory molecules are DNA helper strands that bind to regions flanking the binding site on an initiator.

Furthermore, the composition can comprise a carrier that facilitates the introduction of nucleic acids, such as, for example, nucleic acid monomers and accessory nucleic acid molecules, into a cell, such as a cell containing an initiator associated with a disease or disorder. Carriers for delivery of nucleic acids into cells are well known in the art and examples are described above.

In some embodiments, a computer program is provided that designs and/or aids in the design of the primary sequences of hairpin monomers. In some embodiments, a program can be used that specifies assembly and/or disassembly pathways for dynamic functions using nodal abstractions. In some embodiments, the program translates nodal abstractions into hairpin motifs. In some embodiments, the program designs primary sequences of hairpin monomers. In this manner, primary sequences for hairpin monomers for implement-
ing a dynamic pathway can be provided by the program. In some embodiments, the program performs any of the methods described herein.

Molecular Compilers

Within the nucleic acid design community, it is common practice to specify a design as a set of one or more static target secondary structures. Seeman, Nature 421, 427-431(2003). The sequences of the constituent strands are then typically designed by optimizing an objective function that captures some combination of affinity and/or specificity for the target structures. Seeman, J. Biomol. Struct. Dyn. 8, 573-581 (1990); Hofacker et al., Chem. Mon. 125, 167-188 (1994); Andronescu et al., J. Mol. Biol. 336, 607-624 (2004); Dirks et al., Nucleic Acids Res. 32, 1392-1403 (2004).

By contrast, dynamic function encoded in a self-assembly system can be designed by programming the reaction pathway of the system as described herein. The intended dynamic function is first specified using a reaction graph. The reaction graph is then implemented in terms of the present hairpin motif, and finally the molecular implementation is encoded in the primary sequences of a set of nucleic acid strands of the hairpin monomers. As such, the standardized hairpin motif and the reaction graph provide layers of abstraction that bridge the description of the dynamic behavior of the system and the set of nucleic acid primary sequences, which implement the target behavior.

In some embodiments, automating the process depicted in the reaction graph can provide a biomolecular compiler that can take the desired dynamic function as input, translate it first to a reaction graph, then to a motif-based molecular implementation, and subsequently into nucleic acid sequences that encode the intended dynamic function.

In some embodiments, a method for preparing hairpin monomers for carrying out a dynamic function is provided. In some embodiments, the method includes: providing an input, switching the state of a first input port on a first node abstraction from accessible to inaccessible, switching the state of a second input port on a second node abstraction from accessible to inaccessible and switching the state of the first output port on the first node abstraction from accessible to inaccessible; and designing a first hairpin monomer based on the first node abstraction and a second hairpin monomer based on the second node abstraction, wherein the first and second hairpin monomers self-assemble in the presence of an initiator to perform the dynamic function. In some embodiments, a step of designing nucleic acid primary sequences for the first hairpin monomer and second hairpin monomer can be included in the method.

EXAMPLES

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

Example 1

System Synthesis

This example illustrates the synthesis and preparation of hairpin monomers.

Nucleic Acid Synthesis

Nucleic acid, such as, for example, DNA, can be synthesized and purified by various methods known in the art. In some embodiments, purified nucleic acid strands can be reconstituted in, for example, ultrapure water with resistance of at least 18 MΩ. The concentrations of the nucleic acid solutions can then be determined by measurement of ultraviolet absorption at 260 nm.

Hairpin Synthesis

Each hairpin can be synthesized using a variety of methods known in the art. For example, in some embodiments, two nucleic acid pieces can be synthesized and ligated to produce the full hairpin. The ligation can be performed using a suitable enzyme, such as, for example, T4 DNA ligase at suitable conditions. For example, the ligation can be performed at room temperature or 16°C for at least two hours. Ligated strands can then be further purified using, for example, denaturing polyacrylamide gel electrophoresis. The bands corresponding to the nucleic acid strands of expected sizes can be visualized by, for example, UV shadowing and excised from the gel. The nucleic acid strands can then be eluted and recovered by ethanol precipitation.

Monomer Preparation

For preparation of monomers, concentrated nucleic acid strands can be diluted to suitable reaction conditions. Exemplary reaction conditions can be as follows: 50 mM NaH2PO4, 0.5 M NaCl, pH 6.8; or 20 mM Tris, pH 7.6, 2 mM EDTA, 12.5 mM Mg2+ (1×TAE/Mg2+ buffer). The hairpins are allowed to anneal under suitable conditions. For example, the hairpins can then annealed by, for example, heating for 5 minutes at 90°C, and then turning off the heating block to allow the system to cool to room temperature (requiring at least 2 hours).

Example 2

System Design Example: Catalytic Formation of 3-Arm Junction

This example illustrates the design procedure for the catalytic 3-arm junction system as presented in FIGS. 2a and b. Step (1), Pathway Specification. The desired dynamic behavior (FIG. 6a) is specified using a reaction graph (FIG. 6b).

Step (2.1), Basic Molecular Logic Implementation. The reaction graph is implemented using the standard motif (FIG. 6c).

Step (2.2),Padding/Clamping. The basic implementation (FIG. 6c) is modified by adding padding/clamping segments, i.e., segments x, y, z, x*, y*, and z* in FIG. 6d. These segments serve two purposes. First, they serve as ‘padding’ segments to modulate the lengths of the hairpin’s exposed toehold, stem, and loop, which permits more flexible dimensioning in the next step. Second, the segments serve as ‘clamps’ to decrease spurious ‘leakage’ reactions in the absence of the initiators. Consider un-clamped hairpin A and hairpin B in FIG. 6c. When the left-end of hairpin A’s stem ‘breathes,’ the 3’ end of the segment b* will be transiently exposed, revealing a partial toehold that is complementary to the toehold b of hairpin B. This transient toehold exposure would permit hairpin A and hairpin B to reach spuriously and form A·B (which would then react with C to form A·B·C). By contrast, the ‘breathing’ of the left end of the clamped hairpin A stem in Figure 8a/d exposes x* instead of b* remains sequestered, discouraging spurious reaction between A and B that nucleates at b*.

Step (2.3), Segment Dimensioning. The purpose of segment dimensioning is to assign the length of each segment in terms of the number of nucleotides such that under specified conditions, the desired reaction can proceed smoothly while spurious reactions are suppressed.
The NUPACK server (www.nupack.org) can be used for dimensioning. For the catalytic 3-arm junction system described here, assigning 6 nt to each segment (Fig. 6e) stabilizes critical structures in the reaction pathway in the context of a dilute solution of interacting nucleic acid strands.

Step (3), Sequence Design.

Based on the criteria determined from implementing the reaction graph, primary sequences for the hairpin monomers are designed. The sequences are optimized computationally to maximize affinity and specificity for the formation of the 3-arm junction by minimizing the average number of incorrectly paired bases at equilibrium. The system is synthesized as described in Example 1. Verification of the system is carried out using gel and single-molecule AFM.

Example 3

Execution of the Reaction Graphs for Catalytic 3-Arm/4-Arm Junction Systems

This example illustrates the step-by-step execution of the reaction graphs in Figs. 2a and c. As shown in Fig. 7a, Reaction 1 (assembly), a bond is made between the accessible output port of A and the accessible input port of B and both ports are flipped to inaccessible states; the output port of A is flipped to the accessible state (based on the internal logic of node A). Reaction 2 (assembly): a bond is made between the newly accessible output port of A and the accessible input port of B and both ports are flipped to inaccessible states; the output port of B is flipped to the accessible state (based on the internal logic of node B). Reaction 3 (assembly): A bond is made between the newly accessible output port of B and the input port of C and both ports are flipped to inaccessible states; the output port of C is flipped to the accessible state (based on the internal logic of node C). Reaction 4 (disassembly): the bond between the inaccessible output port of I and the inaccessible input port of A is displaced by a bond between the newly accessible blue output port of C and the input port of A; the states of the two output ports are flipped.

The reaction graph in Fig. 7a contains a k=3 disassembly cycle: input port of A o blue output of A o input port of B o blue output port of B o input port of C o blue output port of C o input port of A. The reaction graph in Fig. 7b contains a k=4 disassembly cycle: input port of A o blue output of A o input port of B o blue output port of B o input port of C o blue output port of C o input port of D o blue output port of D o input port of A.

Example 4

Catalytic Structure Formation: Catalytic Formation of a 4-Arm Junction

Figs. 8a and b depict the reaction graph and reaction schematic for the catalytic formation of a 4-arm junction, respectively. In the absence of initiator I, hairpins A, B, C, and D are kinetically impeded from forming the 4-arm junction that is predicted to dominate at equilibrium. Introduction of I into the system (Fig. 8b, bottom) activates a cascade of assembly steps with A, B, C, and D followed by a disassembly step in which D displaces I from the complex, freeing I to catalyze the self-assembly of additional branched junctions. The lengths of segments q, r, s, t, and r are 18 nt; the lengths of the other segments are 6 nt. Hairpins A, B, C, and D are metastable in the absence of the initiator I. The initiator I catalyzes monomers A, B, C, and D to form a 4-arm DNA junction, as follows: (1) segment a of I nucleates at the toehold of hairpin A and initiates a strand displacement that results in the opening of hairpin A; (2) newly exposed b of A nucleates at toehold of B and results in the opening of B; (3) newly exposed c of B nucleates at toehold of C and results in the opening of C; (4) newly exposed d of C nucleates at d of hairpin D and results in the opening of D; (4b) D displaces I from A.

Lanes 1-5: A gel shifting assay validates each reaction step depicted in panel (b). Lanes 5-9: Effects of different concentrations of I (1, 0.5x, 0.25x, 0.1x, and 0x) on the formation of A•B•C•D. 600 nM reactants were incubated at room temperature for 2 hours. Lane 10: A•B•C•D annealed over 2.5 hours (600 nM hairpin species heated at 95°C for 5 minutes and cooled to room temperature over 2.5 hrs). The 2% agarose gel was prepared in 1xLB buffer (Faster Better Media, LLC) with 0.5 µg/ml ethidium bromide. The gels were run at 150 V for 30 min at room temperature and then visualized using UV transillumination. The hairpins used for these reactions did not contain the 3' tails (p, r, s, t, and r).

Native agarose gel electrophoresis (Fig. 8c) confirms that the hairpins assemble slowly in the absence of the initiator (Lane 9) and that assembly is dramatically accelerated by the addition of initiator (Lane 5). Disassembly of the initiator enables catalytic turnover as indicated by the nearly complete consumption of hairpins even at sub-stoichiometric initiator concentrations (Lanes 6-8). As in the 3-arm junction case, only minimal assembly is achieved by annealing the hairpin mixture (Lane 10).

AFM imaging of the catalyzed self-assembly product (augmented with strands that extend the duplex portion of each arm as described in the caption) reveals the expected 4-arm junction morphology (Fig. 8d). To assist in AFM imaging of the 4-arm junction, four strands (Ae, Bc, Ce, and De) were incubated with the catalytically formed 4-arm junction A•B•C•D. Note that the duplex portion of the arms of the final structure A•B•C•D•Ae•Bc•Ce•De are twice as long as the duplex portion of the arms of A•B•C•D. Two AFM images of A•B•C•D•Ae•Bc•Ce•De are presented in Fig. 8d.

Example 5

AFM Image Analysis

Figs. 9a and b depict AFM image analysis of 3-arm/4-arm junctions. Using a B-DNA model where one helical turn contains 10.5 base pairs and measures 3.4 nm, the expected arm length for the 3-arm junction was calculated as follows: (24/10.5)x3.4 nm=7.8 nm. Similarly, the arm length for the 4-arm junction is calculated to be 7.8+7.8=15.6 nm. The measured lengths of the arms are roughly consistent with the calculated lengths. Figs. 10a and b show AFM images with a larger field of view for 3-arm (a) and 4-arm (b) junctions.

Example 6

Design for the Catalytic Formation of a k-Arm Junction

The catalytic system described in Fig. 2 and Fig. 8 can, in principle, be generalized to a system capable of the catalytic formation of a k-arm junction. Fig. 11a and b describe the reaction graph and the secondary structure schematic for the catalytic formation of a k-arm junction. Hairpins H1, H2, . . . , Hk are metastable in the absence of the initiator I. The initiator I catalyzes monomers H1, H2, . . . , Hk to form a k-arm DNA junction.
FIGS. 12a-c depict an example when k=6. In FIGS. 12a-c, hairpins H1, H2, H3, H4, H5, and H6 are metastable in the absence of the initiator I. The initiator I catalyzes monomers H1, H2, H3, H4, H5, and H6 to form a six-arm DNA junction as follows:

Step 1: segment α of I nucleates at the toehold α of hairpin H1 and initiates a strand displacement that results in the opening of hairpin H1.

Step 2: the newly exposed aβ of H1 nucleates at the toehold aβ of hairpin H2 and opens hairpin H2.

Step 3: the newly exposed aγ of H2 nucleates at the toehold aγ of hairpin H3 and opens hairpin H3.

Step 4: the newly exposed aδ of H3 nucleates at the toehold aδ of hairpin H4 and opens hairpin H4.

Step 5: the newly exposed aε of H4 nucleates at the toehold aε of hairpin H5, and opens hairpin H5.

Step 6: the newly exposed aζ of H5 nucleates at the toehold aζ of hairpin H6, and opens hairpin H6.

Step 7: H6 displaces I from H1.

Example 7

Catalytic Circuitry

FIG. 13 describes the step-by-step execution of the reaction in FIG. 3a. The reaction starts at solid arrow (1) that connects the accessible output port of I and the accessible input port of A. Note that by convention, the two arrows entering the same input port of A depict parallel processes on separate copies of the nodal species.

Reaction 1 (Assembly):

- A bond is made between the accessible output port of I and the accessible input port of A and both ports are flipped to inaccessible states; the output port of A is flipped to the accessible state (based on the internal logic of node A).

Reaction 2a (Assembly):

- A bond is made between the newly accessible output port of A and the accessible input port of B and both ports are flipped to inaccessible states; the two output ports of B are flipped to accessible states (based on the internal logic of node B).

Reaction 2b (Disassembly):

- The bond between the inaccessible output port of I and the inaccessible input port of A is displaced by a bond between the newly accessible blue output port of B and the input port of A; the states of the two output ports are flipped.

Reaction 3 (Assembly):

- A bond is made between the newly accessible green output port of B and the accessible input port of C and both ports are flipped to inaccessible states; the output port of C is flipped to the accessible state (based on the internal logic of node C).

Reaction 4a (Assembly):

- A bond is made between the newly accessible output port of C and the accessible input port of D and both ports are flipped to inaccessible states; the output ports of D are flipped to accessible states (based on the internal logic of node D).

Reaction 4b (Disassembly):

- The bond between the inaccessible green output port of B and the inaccessible input port of C is displaced by a bond between the newly accessible blue output port of D and the input port of C; the states of the two output ports are flipped.

Reaction 5 (Assembly):

- A bond is made between the newly accessible green output port of D and the accessible input port of A and both ports are flipped to inaccessible states; the output port of A is flipped to the accessible state (based on the internal logic of node A).

Reaction 6a (Assembly):

- A bond is made between the newly accessible output port of A and the accessible input port of B and both ports are flipped to inaccessible states; the output ports of B are flipped to accessible states (based on the internal logic of node B).

Reaction 6b (Disassembly):

- The bond between the inaccessible green output port of D and the inaccessible input port of A is displaced by a bond between the newly accessible blue output port of B and the input port of A; the states of the two output ports are flipped.

FIG. 14 describes the detailed reaction flow of the autocatalytic system described in FIG. 3. FIGS. 15a-c describe additional intermediate steps. Steps 1-2 are the initiation stage; steps 3-6 are the exponential amplification stage.

Step 1: the toehold a of I nucleates at the toehold a of A, resulting in the opening of the hairpin and the formation of the product A•I.

Step 2: A•I, with b newlly exposed, opens hairpin B (step 2a); B subsequently displaces I from A (step 2b), producing A•B and bringing the system to the exponential amplification stage. The single-stranded tail (ν,ν-ν + α,α* + ε,ε) of A•B next catalyzes C and D to form C•D (in steps 3 and 4).

Step 3: A•B, with c newly exposed, opens hairpin C.

Step 4: A•B•C, with d newly exposed, opens hairpin D (step 4a); D subsequently displaces C from B, separating A•B and C•D (step 4b). The single-stranded tail (ν,ν,ν + α,α* + ε,ε) of C•D is identical to I and next catalyzes A and B to form A•B (in steps 5 and 6).

Step 5: C•D, with a newly exposed, opens hairpin A.

Step 6: C•D•A, with b newly exposed, opens B (step 6a); B subsequently displaces A from D, separating C•D and A•B (step 6b).

FIGS. 16a and b depict a stepping gel for the autocatalytic system. The hairpins used for these reactions were synthesized and purified by IDT DNA and used without further purification. The annealed samples were annealed at 2 μM reaction concentrations: heating at 95°C for 5 minutes followed by cooling to room temperature over approximately 2.5 hours. The room temperature reactions were conducted with each reactant species at 1 μM concentration. Consider the sample, (A1)+B, in Lane 5. The sample was prepared by first annealing a mix containing 2 μM A and 2 μM I to produce (A1). Then 2 μL of (A1), at 2 μM concentration, was mixed with 2 μL of B at 2 μM concentration and allowed to react at room temperature for 15 minutes. Lanes 1 and 14 are 20-1000 bp DNA ladders (Bio-Rad). The 5% native polyacrylamide gel was prepared in 1×TAE/Mg++ buffer (20 mM Tris, pH=7.6, 2 mM EDTA, 12.5 mM Mg++). The samples were loaded with 10% glycerol. The gel was run at 100 V for 90 minutes at room temperature, post-stained with 0.5 μg/mL ethidium bromide, and visualized by UV transillumination.

The blue line delineates the boundary between two gels.

The autocatalytic system was validated on a step-by-step basis using native polyacrylamide gel electrophoresis (PAGE) (FIG. 16b):

Step 1: Hairpin A reacts with initiator I and produces a band that corresponds to product A•I (Lane 3), which migrates at about the same speed as the annealed product A•I (Lane 4), as expected.

Step 2. Annealed sample A•I reacts with hairpin B and produces a band that corresponds to product A•B (Lane 5), which migrates at about the same speed as the annealed product A•B (Lane 6), as expected.

Step 3. Annealed sample A•B reacts with hairpin C and produces a band that corresponds to product A•B•C (Lane 7), which migrates at about the same speed as the annealed product A•B•C (Lane 8), as expected.

Step 4. Annealed sample A•B•C reacts with hairpin D and produces a band that corresponds to product A•B•D and C•D.
(Lane 9), which migrates at about the same speed as the annealed product A-B (Lane 6) and the annealed product C-D (Lane 10), as expected.

Step 5. Annealed sample C-D reacts with hairpin A and produces a band that corresponds to product C-D:A (Lane 11), which migrates at about the same speed as the annealed product C-D:A (Lane 12), as expected.

Step 6. Annealed sample C-D:A reacts with hairpin B and produces a band that corresponds to product C-D and A-B (Lane 13), which migrates at about the same speed as the annealed product C-D (Lane 10) and the annealed product A-B (Lane 6), as expected.

System kinetic analysis and data analysis is described in Yin et al., *Nature* 451(7176), 318-322; Supplementary Information pages 1-49 (2008), which is incorporated herein by reference in its entirety.

Example 8

Nucleated Dendritic Growth

FIG. 17 depicts the execution of the reaction graph of FIG. 4a. The multiple arrows entering the same input port depict parallel processes on separate copies of the nodal species. The parallel processes are not synchronized and hence it is possible, for example, that after A1 assembles with A2, the assembly of A2 with A3 occurs before the assembly of A1 with B2.

FIG. 18 and FIG. 19 present the detailed reaction schematic of the nucleated dendritic growth system described in FIG. 4. In the absence of the inhibitor I, hairpin monomers co-exist metastably. The initiator I triggers the system to self-assemble into a binary tree of a prescribed size.

Step 1: The toehold a* of the initiator I nucleates at the toehold a of hairpin A1, resulting in the opening of A1 and the formation of the first generation dendrimer, G1.

Step 2: A1, with b* and c* newly exposed, opens hairpins A2 and B2, producing the second generation dendrimer, G2. Note that now A2 and B2 reveal single-stranded tails of identical sequences.

Step 3: A2 and B2, with d* and e* newly exposed, open hairpins A3 and B3, producing G3.

Step 4: Each copy of A3 and B3, with its newly exposed f* and g*, opens hairpins A4 and B4, producing G4.

Step 5: Each copy of A4 and B4, with its newly exposed h* and i*, opens hairpins A5 and B5, producing G5.

Example 9

Quantitative Amplification Gel and AFM Image Analysis

FIG. 20 demonstrates that the concentration of dendrimer depends linearly on the concentration of the initiator in the system. The top panel of FIG. 20 shows different concentrations of initiator incubated with all hairpin species (A1, A2, B2, 91 nM; the concentration doubles for each subsequent generation of hairpins). The gel shown in FIG. 20 is used to measure fluorescence emission from Cy5, which is used to label hairpin A1. In the figure, D denotes dendrimers; M denotes monomers. The bottom panel of FIG. 20 shows linear fit between the fluorescence signal of the dominant reaction product versus the concentration of initiator. Data from three independent experiments are denoted respectively by blue crosses, red diamonds, and green circles. Each set of data is normalized by the signal obtained at 70 nM initiator concentration.

Using a B-DNA model where one helical turn contains 10.5 base pairs and measures 3.4 nm, we calculate the expected arm length for the duplex formed by A1 and I to be 25/10.5x3.4 nm=8.1 nm and the approximate length of all the other duplex segments to be 50/10.5x3.4 nm=16.2 nm. FIG. 21 shows the image analysis for G3 and G4 dendrimers. The small images are screenshots of the measurement section files. The distance between the two red arrowheads is listed above the image. The blue arrow points to the 4-arm junction in both the schematic and the images and help to relate the images to the schematic. FIG. 22 shows the image analysis for G5 dendrimers. The measured lengths of the arms are roughly consistent with the calculated lengths. In FIG. 22, the distance between the two red arrowheads is listed above the image. The blue arrow points to the 5-arm/4-arm junctions in both the schematic and the images and help to relate the images to the schematic. Note that the duplex in the left panel, likely due to damage during sample preparation or AFM scanning. FIG. 23 shows a large field-of-view AFM image of the G5 system. As seen by FIGS. 21-23, in most AFM images, only the duplex portions of the dendrimer are visible.

Example 10

Autonomous Locomotion

FIGS. 24a-b depicts the step-by-step execution of the reaction graph for the walker. In FIGS. 24a-b, the reaction steps corresponding to the progressive sub-population of walkers are shown in purple. In the initial conditions prior to Step 1, the input ports of the A nodes at sites 1 and 2 are bound to the output ports of the 1 nodes on the bipedal walker. Execution begins with an assembly reaction between the accessible output port on either of these A nodes and the accessible input port on B. In the top route of Step 1, B assemblies with A at site 1, resulting in the disassembly of the trailing 1 from A, which is then free to assemble with A at site 3, moving the walker one step down the track and bringing the system to Step 2. Alternatively, a B node could bind to A at site 2 prior to the assembly of I with A at site 3, resulting in the disassembly of the walker from the track. The walker could then diffuse through the bulk solution and re-attach to the same track or another track at any A monomer that has not yet been occupied. In the bottom route of Step 1, node B assemblies with node A at site 2, resulting in the disassembly of the leading 1 from A. Due to geometric constraints (inextensible walker torso and rigid track backbone), the walker cannot attach to site 1 and site 3 simultaneously and hence will eventually detach from the track when a B node assembles with A at site 1. Similarly, in Step 2 and Step 3, progressive stepping occurs stochastically for a sub-population of walkers. In Step 4, the walker will disassemble from track.

According to FIGS. 24a-b, the initial bond between the output port of I and the input port of A indicates that an assembly reaction has already occurred prior to the execution of the reaction graph. As noted above, static structural elements can impose geometrical constraints on the execution of the reaction graph. In the reaction graph depicted here, the gray structural elements represent a rigid track backbone and an inextensible walker torso; their relative dimensions imply that when one I node is attached to an A node on the track, the other I node can only interact with the A node to either side. A secondary structure schematic of the walker system of FIG. 5 is shown in FIG. 25. FIGS. 26a-b depict the step-by-
The walker system was assembled in four steps (FIG. 28a).

Step 0. The walker (W) was assembled by annealing strands W1-BHQ1 and W2-BHQ1 as follows: heat the mixture at 95°C for 5 minutes and slowly cool to room temperature over the course of 4 hours.

Step 1. Hairpins S1 and S4 were mixed with track strands S2, S3, and S8, then annealed to produce Track 1 (T1) as follows: heat the system at 95°C for 5 minutes; slowly cool to room temperature over the course of 4 hours.

Step 2. T1 and the pre-assembled walker (W) were incubated at room temperature for 2 hours to produce T1+W.

Step 3. Hairpins S6, S9, and S11 were mixed with track strands S7, S8, S10, and S12, then annealed to produce Track 2 (T2) as follows: heat the system to 95°C for 5 minutes; slowly cool to room temperature over the course of 4 hours. For the bidental and monopodal landing control experiments, the S7 track strand is replaced by S7 truncated so that T1 and T2 remain disjoint.

Step 4. T2 and T1+W were incubated at room temperature for 3 hours to produce the final system, T14+W×T2.

Native agarose gel electrophoresis demonstrates a band shifting pattern that confirms on a step-by-step basis the correct assembly of the walker system. (FIG. 28b). Samples were annealed and assembled in reaction buffer (4 mM MgCl₂, 15 mM KCl, and 10 mM Tris-HCl, pH=8.0) with all species at 0.5 μM. A 3% native agarose gel was prepared in 1xLB buffer (Faster Better Media, LLC). Samples were loaded with 2xSYBR Gold stain (Invitrogen) and 10% glycerol. The gel was run at 200 V for 40 minutes at room temperature and visualized using an FLA-5100 imaging system (Fuji Photo Film Co., Ltd.).

Example 11

Characterization of the Fuel System

This example describes the fuel system that powers the walker system, which is depicted in FIGS. 29a and b. Here, hairpins A and B co-exist metastably in the absence of catalyst I. Catalyst I catalyzes A and B to form duplex AB (FIG. 29a). Native gel electrophoresis (FIG. 29b) confirms that the hairpins assemble slowly in the absence of the initiator (Lane 7) and that the assembly is dramatically accelerated by the addition of initiator (Lane 3). Disassembly of the initiator enables catalytic turnover as indicated by the nearly complete consumption of hairpins even at sub-stoichiometric initiator concentrations (Lanes 4-6).

Agarose gel electrophoresis demonstrates catalytic formation of the DNA duplex (FIG. 29b). The hairpins were prepared in reaction buffer (4 mM MgCib, 15 mM KCl, and 10 mM Tris-HCl, pH=8.0) using a snap-cooling procedure: heating at 90°C for 5 minutes and cooling on ice for 1 minute. The hairpins were allowed to equilibrate at room temperature for 30 minutes before use. Lanes 1-3: A gel shifting assay verifies each reaction step depicted in panel (a). Lanes 3-7: Effects of different concentrations of T (1x, 0.5x, 0.25x, 0.1x, and 0x) on the formation of A+B. Reactants were incubated at 1 μM at room temperature for 2 hours. Lane 8: A+B annealed over 2.5 hours (1 μM hairpin species heated at 95°C C for 5 minutes and cooled to room temperature over 2.5 hrs). Upon completion of the reaction, the samples were loaded with 5xSYBR Gold stain (Invitrogen) and 10% glycerol into a 2% native agarose gel, prepared with 1xLB buffer (Faster Better Media, LLC). The gel was run at 350 V for 10 minutes at room temperature and visualized using an FLA-5100 imaging system (Fuji Photo Film Co., Ltd.).

Fluorescence quenching experiments was carried out to investigate catalyst recovery, and is described in Yin et al., Nature 451(7176), 318-322; Supplementary Information pages 1-49 (2008) at page 31-37, which is incorporated herein by reference in its entirety.

Example 13

Synthesis of DNA, Hairpins and Monomers

This example illustrates the synthesis of DNA, hairpins and monomers used in the Examples.

DNA was synthesized and purified by Integrated DNA Technologies. The purified DNA strands were reconstituted in ultrapure water (resistance of 18 MQ2-cm). The concentrations of the DNA solutions were determined by the measurement of ultraviolet light absorption at 260 nm.

Each hairpin was synthesized as two pieces which were then ligated to produce the full hairpin. The ligation was performed using T4 DNA ligase (New England Biolabs) at either room temperature or 16°C. for a minimum of two hours. Ligated strands were further purified using denaturing polyacrylamide gel electrophoresis. The bands corresponding to the DNA strands of expected sizes were visualized by UV shadowing and excised from the gel. The DNA strands were then eluted and recovered by ethanol precipitation. (3) Monomer preparation. The concentrated DNA strands were diluted to reaction conditions: 50 mM Na₂HPO₄, 0.5 M NaCl, pH=6.8 for species in FIG. 2 and FIG. S4; and 20 mM Tris, pH=7.6, 2 mM EDTA, 12.5 mM Mg** (1xTAE/Mg** buffer) for species in FIG. 3, FIG. S12, and FIG. 4. The hairpins were then annealed by heating for 5 minutes at 90°C, and then the heating block was turned off to allow the system to cool to room temperature (requiring at least 2 hours).

Commercially available synthetic single-stranded DNA oligos can be impure and contain incorrectly synthesized strands. The presence of such erroneous strands can contribute to leakage during self-assembly. In some embodiments, to improve strand purity and hence decrease system leakage, the following enzyme-based ligation method can be used to synthesize the hairpin monomers: two constituent segments of a hairpin are synthesized and purified separately and ligated to produce the full hairpin (FIG. 30). Significant reduction of system leakage in the ligation-based system is observed, as compared to the un-ligated system.

The observed error reduction may be attributed to the following two mechanisms. First, longer DNA strands contain more errors than shorter fragments, since the shorter fragments can be purified to a higher purity. Behlke et al., Tech. Rep., Integrated DNA Technologies (2005). As such, the two constituent segments contain fewer total errors than the full strand. Second, for T4 ligase mediated ligation of Hα and Hβ, the successful ligation depends on the correct juxtaposition of the 5’ end of fragment Hβ with the 3’ end of fragment Hα. This requirement provides an additional error reduction mechanism: DNA segments with errors in the regions adjacent to the nick position are not ligated successfully and are eliminated during the subsequent gel purification.
Example 14

Gel Electrophoresis

This example illustrates electrophoresis analysis of the self-assembly systems. For the gel in FIG. 2c, 12 μL of each 3 μM hairpin species were mixed by pipetting. 6 μL of this master mix was aliquoted into 5 separate tubes. To these tubes were added 2 μL of either 3 μM (Lane 1), 1.5 μM (Lane 2), 0.75 μM (Lane 3), 0.3 μM (Lane 4), or 1× reaction buffer (50 mM Na-HPO₄, 0.5 M NaCl, pH=6.8) (Lane 5) to reach a total reaction volume of 8 μL. The samples were then mixed by pipetting and allowed to react for 2.5 hours at room temperature. The annealed reaction (Lane 6), prepared 0.5 hour in advance, was made by mixing 2 μL of each hairpin with 2 μL of the 1× reaction buffer, and then annealing as described in monomer preparation. A 2% native agarose gel was prepared for use in 1× TBE buffer (Faster Better Media, LLC). 1 μL of each sample was then mixed with 1 μL of 5×SYBR Gold loading buffer: 50% glyceral/50% H₂O/SYBR Gold (Invitrogen) and loaded into the gel. The gel was run at 350 V for 10 minutes at room temperature and imaged using an FLA-5100 imaging system (Fujifilm Photo Film Co., Ltd.). For the gel in FIG. 4c, hairpins were annealed at the following concentrations: A₁, A₂, B₂, A₃, and B₃ at 1 μM; A₄ and B₄ at 2 μM; A₅ and B₅ at 4 μM. The initiator I was prepared at 800 nM. The following sample mixtures were prepared: Lane 1: A₁, Lane 2: A₁+I; Lane 3: A₁+I+A₁+B₂; Lane 4: A₁+I+A₁+B₂+A₁+B₃; Lane 5: A₁+I+A₁+B₂+A₁+B₃+A₂+B₄; Lane 6: A₁+I+A₁+B₂+A₁+B₃+A₂+B₄+A₅+B₅; Lane 7: A₁+I+A₁+B₂+A₁+B₃+A₂+B₄+A₅+B₆; Lane 8: A₁+I+A₁+B₂+A₁+B₃+A₂+B₄+A₅+B₆. The samples were mixed by pipetting and allowed to react for 1 hour at room temperature. The sample was then mixed with 5×LB loading buffer (Faster Better Media, LLC) to reach 1× loading buffer concentration (8 μL sample, 2 μL loading buffer). The sample/loading buffer mix was loaded into a 1% native agarose gel prepared in 1×LB buffer. The gel was run at 350 V for 10 minutes at room temperature and then imaged and quantified using an FLA-5100 imaging system. The reactions in FIG. 3d, the hairpins were mixed to reach the following final concentrations: A₁-C₅, A₂, B₂, 100 nM; A₃, B₃, 200 nM; A₄, B₄, 400 nM; A₅, B₅, 800 nM. 9 μL of this mix were then aliquoted into 10 separate tubes. To these tubes was added either 1×TAE/Mg²⁺ reaction buffer or the initiator I to give the indicated final concentration of I and a final volume of 11 μL. The samples were mixed by pipetting and then allowed to react for 1 hour at room temperature. The sample was then mixed with 5×LB loading buffer (Faster Better Media, LLC) to reach 1× loading buffer concentration (8 μL sample, 2 μL loading buffer). The sample/loading buffer mix was loaded into a 1% native agarose gel prepared in 1×LB buffer. The gel was run at 350 V for 10 minutes at room temperature and then imaged and quantified using an FLA-5100 imaging system. The experiments were performed with 10 μM inert 25-nt poly-T carrier strands in the reaction solution.

Example 15

AFM Imaging

This example illustrates AFM imaging of the self-assembly systems. AFM images were obtained using a multimode scanning probe microscope (Veeco Instruments Inc.), equipped with a Q-Control module for analog AFM systems (Atomic Force F&E GmbH). The images were obtained in liquid phase under tapping mode using DNP-S oxide sharpened silicon nitride cantilevers (Veeco Instruments Inc.). Samples were first diluted in 1×TAE/Mg²⁺ buffer to achieve the desired imaging density. A 20 μL drop of 1×TAE/Mg²⁺ and a 5 μL drop of sample were applied to the surface of freshly cleaved mica and allowed to bind for approximately 2 minutes. Supplemental 15-30 mM Ni²⁺ was added to increase the strength of DNA-mica binding. Before placing the fluid cell on top of the mica puck, an additional 15-20 μL of 1×TAE/Mg²⁺ buffer was added to the cavity between the fluid cell and the AFM cantilever chip to avoid bubbles.

Example 16

Fluorescence Experiments

This example illustrates fluorescence experiments with the self-assembly systems. Catalytic circuity. Fluorescence data were obtained using a QM-6/2005 steady state spectrofluorometer (Photon Technology International), equipped with a Turret 400™ four-position cuvette holder (Quantum Northwest) and 3.5 mL quartz cuvettes (Hellma GmbH & Co. KG). The temperature was set to 25°C. The excitation and emission wavelengths were set to 520 nm (2 nm bandwidth) and 540 nm (4 nm bandwidth), respectively. For the experiments in FIG. 3c, hairpin monomers, A, B, C, and D, and initiator, I, were prepared separately as described above. A 40 μL 1 μM A was added to 1800 μL 1×TAE/Mg²⁺ buffer and mixed by rapid pipetting 8 times using a 1 ml tip. The baseline signal was recorded for 16 minutes. Then 40 μL of 1 μM B, C, and D, and the appropriate concentration of I (or 1×TAE/Mg²⁺ buffer in the case of 0μl) were added to the cuvette (to reach the target concentrations described in FIG. 3c) and mixed by rapid pipetting 8 times using a 1 ml tip. The control with 20 nM A alone was monitored continuously. The final volume was 2 mL for all experiments. The experiments were performed with 10 μM inert 25-nt poly-T carrier strand in the individual hairpin and initiator stock solutions and ~1 μM inert 25-nt poly-T carrier strands in the final reaction solution.

Autonomous locomotion. Fluorescence experiments were performed at 21°C using the same spectrofluorometer as above. Two 3.5 mL quartz cuvettes (Hellma GmbH & Co. KG) were used in each set of experiments. Excitation and emission wavelengths were set to 492 nm and 517 nm (for FAM), 527 nm and 551 nm (for JOE), and 558 nm and 578 nm (for TAMRA), respectively, with 4 nm bandwidths. The assembly of the walker system is described above. Hairpin B was snap cooled in the reaction buffer (4 mM MgCl₂, 15 mM KCl, and 10 mM Tris-HCl, pH=8.0) heating at 95°C for 90 seconds, rapid cooling at room temperature, allowed to sit at room temperature for 30 minutes before use. The system was assembled using 4 nM track and 3.5 nM bipedal walker. A sub-stoichiometric amount of walker was used to ensure that no free-floating walker would bind to hairpin A on the track. For the same reason, sub-stoichiometric monopedal walker (7 nM) was used in the diffusion experiments. The final concentration of hairpin B was 20 nM, which was equimolar with the five A hairpins on the track (5×4 nM=20 nM). The assembled track was first introduced to record the fluorescence baselines for FAM, JOE, and TAMRA. Hairpin B was then introduced and mixed 100 times by rapid pipetting to start walker locomotion.

Example 17

DNA Sequences

The DNA sequences for the systems described in the Examples are presented both as secondary structure schematics in FIGS. 31-35 and as text sequences annotated with segment names.
For each hairpin sequence X, the two segments that are ligated to produce X are indicated as Xa and Xb. Strand modifications are indicated as follows:

3' phosphorylation: /5'Phos;
3'6-carboxyfluorescein: /36FAM/;
5'6-carboxyfluorescein: /56FAM/;
5'6-carboxy-4',5' dichloro-2'-7'-dimethoxyfluorescein (NHIS Ester): /5JOE/;
5' tetraclorofluorescein: /5STET/;
5' carboxytetramethylrhodamine (NHIS Ester): /5TMRN/;
3' black hole quencher-1: /3BHQ-1/.

Catalytic 3-arm junction system. The sequences are listed below as text sequences annotated with segment names.

A: a-x-b-y-z-c-y-x*b*x*

GCTGA-AIGTGG-AGGGAG-TAGTCG-TCAAT-CAACA-CGCACT-
CTCCCC-ACACT

SEQ ID NO: 1

Ab:

GCTTGAGAIGTGG

SEQ ID NO: 2

/5Phos/GAGTGGTCTCCCATCACAACGCACATCTCCCCAACAT

B: b-y-c-z-x*a*-z*-c*t-y*

AGGGAGTAGTCG-TGTGTC-ATTGGA-ACACTC-TCAAGC-TCCAAT

CACAAC-GCACTA

SEQ ID NO: 4

Ba:

AGGGAGTAGTCGTT

SEQ ID NO: 5

/5Phos/GAGTAGGACATCAACACACACTCTCCCCACACT

Bb:

/5Phos/GTAGATGGACACATCTCAAGCTCACAACACACT

C: c-z-a-x-y*-b*-x*-a*-z*

GTTGTC-ATTGGA-GCTTGAGAIGTGG-CAAGTGTCTCCCCACACT

SEQ ID NO: 7

TCAAGC-TCCAAT

Ca:

GTTGTAAGTCAGCT

SEQ ID NO: 8

/5Phos/TTAGGTACCTCCATCTCCTACACACACT

I: y*b*x*-a*

GCGACTA-CTCCCCT-ACACTAC-TCAAGC

SEQ ID NO: 10

Catalytic 4-arm junction system. The sequences are listed below as text sequences annotated with segment names.

A: a-w-b-x-y*-c*-x*b*-w*-q*

GCTGA-AIGTGG-AGGGAG-TAGTCG-TCAAT-CAACA-GCACTA-
CTCCCC-ACACTC-ACACTC-ACACTC-ACACTC

SEQ ID NO: 11

Ab:

GCTGAGATGGATGAGGATGTTGACTCCACACTCACAACACACT

SEQ ID NO: 12

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 13

Autocatalytic system. The sequences are listed below as text sequences annotated with segment names.

A: x*-v*-b*-y*-u*-c*-a*-x*-y*-b*-v*-x*-a

ACACTCAACAC-GTGAAC-CCACT-CTCACA-ATCTCC-
ACACTAAGTGGATGGACACATCTCAAGCTCACAACACT

SEQ ID NO: 29

/5STET/TT-ACACTGAAACACACCTCCTACACACTCACAACACTCACAACACACT

SEQ ID NO: 30

/5Phos/GACTCCACTACACACTCACAACACACT

SEQ ID NO: 31

/5Phos/GAGTGGACACATCTCAAGCTCACAACACACT

SEQ ID NO: 32

/5Phos/GTGGATGGACACATCTCAAGCTCACAACACACT

SEQ ID NO: 33

/5Phos/TTAGGTACCTCCATCTCCTACACACACT

SEQ ID NO: 34

/5Phos/TTAGGTACCTCCATCTCCTACACACACT

SEQ ID NO: 35

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 36

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 37

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 38

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 39

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 40

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 41

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 42

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 43

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 44

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 45

/5Phos/CTAACCCTCAACACACACACTCACAACACACT

SEQ ID NO: 46
Nucleated dendritic growth system. The sequences are listed below as text sequences annotated with segment names. \(\text{Aib-Cys} \) (together with \(\text{Aib} \)) is used to synthesize \(\text{Cy5} \) labeled hairpin \(\text{A1} \).
Walker system. Sequence B is the same as described above for fuel for the walker system. W1s is used as a split strand for ligating strands W1a and W1b to produce W1; W2s is used as a split strand for ligating strands W2a and W2b to produce W2.

Fuel for the walker system. The sequences are listed below as text sequences annotated with segment names.

A: {text sequence}

B: {text sequence}

C: {text sequence}

D: {text sequence}

E: {text sequence}

F: {text sequence}

G: {text sequence}

H: {text sequence}
The foregoing description and Examples detail certain specific embodiments of the invention and describes the best mode contemplated by the inventors. It will be appreciated, however, that no matter how detailed the foregoing may appear in text, the invention may be practiced in many ways and the invention should be construed in accordance with the appended claims and any equivalents thereof.

While the present teachings have been described in terms of these exemplary embodiments, the skilled artisan will readily understand that numerous variations and modifications of these exemplary embodiments are possible without undue experimentation. All such variations and modifications are within the scope of the current teachings. The foregoing examples are provided to better illustrate the disclosed teachings and are not intended to limit the scope of the teachings presented herein.

All references cited herein, including patents, patent applications, papers, text books, and the like, and the references cited therein, to the extent that they are not already, are hereby incorporated by reference in their entirety. In the event that one or more of the incorporated literature and similar materials differs from or contradicts this application, including but not limited to defined terms, term usage, described techniques, or the like, this application controls.
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 2

gcttgagattg ttagg

15

<210> SEQ ID NO 3
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 3
gagtagagt ccaatcaca cgcaactct cctaaaact

39

<210> SEQ ID NO 4
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 4
aggaggagt ggttgtgat tggaacactc tcaagctca atcacaagct acta

54

<210> SEQ ID NO 5
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 5
aggaggagt gctg

15

<210> SEQ ID NO 6
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 6
gtggattgaa acatccgaag ctccaatcag aacgacagt

39

<210> SEQ ID NO 7
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 7
gtggattgag gacgtgaga tgttcacata tctccatcatt cttaagctc caat

54

<210> SEQ ID NO 8
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 8

gttgattg gagct

<210> SEQ ID NO 9
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 9

tgagatgttg cact actic cc taa catct ca agctic caat

<210> SEQ ID NO 10
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 10

gcactactcc taacatctca aagc

<210> SEQ ID NO 11
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 11

gcttgagatg tgagatgttg cact actic cc taa catct ca agctic caat cacaacgcac tactic cctaa catcaac cac CaCCaCCaCCaCCaCCaCCaCCC

<210> SEQ ID NO 12
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 12

gcttgagatg tgagatgttg cact actic cc taa catct ca agctic caat cacaacgcac tactic cctaa catcaac cac CaCCaCCaCCaCCaCCaCCaCCC

<210> SEQ ID NO 13
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 13

taacatcc cacaacgcac ccaacc

<210> SEQ ID NO 14
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 14

taacatcc cacaacgcac ccaacc
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 14
agggagtgtg ggtgttgtg tggactcat cttacgtcca atcacaagc actaacaaca
  60
cacacaac ac  
  72

<210> SEQ ID NO 15
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 15
agggagtgtg ggtgttgtg tggactcat cttacgtcca atcacc
  45

<210> SEQ ID NO 16
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 16
aacgcactaa caacacac aaaccac
  27

<210> SEQ ID NO 17
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 17
gttgtgtgg gaaggtgtg gagaacactc tcaacactc atctacgttc caatatcctt
  60
ccttctct cc  
  72

<210> SEQ ID NO 18
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<222> LOCATION: 1
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 18
gttgtgtgg gaaggtgtg gagaacactc tcaacactc atctac
  46

<210> SEQ ID NO 19
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 19
cgtccaatat cttcctctcc ctctcc
  26

<210> SEQ ID NO 20
cggtagatga gtgcttgaga tigttgcacta ctic cctaa.ca tot caag cact cattctcit ctit ottctgttct to 60

cggtagatga gtgcttgaga tigttgcacta ctccctaaca tctcaagcac teactctcttct 72

cggtagatga gtgcttgaga tigttgcacta ctccctaaca tctca 46

gcactccttc tctctctcttc tctctt 26

gcactacttc tatactcttc aagc 24

gtggtggttg tagtggtgta tigtt 24

gtggtggttg tagtggtgta tigtt 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 26

gagaagagag gaagagtat tgtg

<210> SEQ ID NO 27
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 27
gagaagagag gaagagtat gagt

<210> SEQ ID NO 28
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 28
aagaagagac aagagtaccc acctccacct tggcaacact ccacaactaa gtgtctcaac
gtgctagtt gtggagat

<210> SEQ ID NO 29
<211> LENGTH: 78
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' tetrachlorofluorescein
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 29
ttccacactga acacgttagg ccatctccat cctggcaacact ccacaactaa aagtgtctcaac
gtgctagtt gtggagat

<210> SEQ ID NO 30
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 18
<223> OTHER INFORMATION: 3' black hole quencher-1
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 30
gtgctagtt gtggagat

<210> SEQ ID NO 31
<211> LENGTH: 96
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 31

gaacacgtg ttcacattcc atctcggcag atctcggcag gttgctggtt 60
agtgggtag attggcggga tggagttggg caggtcagt 96

<210> SEQ ID NO 31
<211> LENGTH: 78
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 32

gaacacgtg ttcacattcc atctcggcag atctcggcag gttgctggtt 60
agtgggtag attggcggga tggagttggg caggtcagt 96

<210> SEQ ID NO 33
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 33

ggtgagagt ggtcagaa 18

<210> SEQ ID NO 34
<211> LENGTH: 78
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 34

tgtgctgctg gtggctgcaat gacgagatcc atcgctgctg aatctcgtcc aacgtgatcc 60
tgctgctgctg ttcacattcc 78

<210> SEQ ID NO 35
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 35

tgtgctgctg gtggctgcaat gacgagatcc atcgctgctg aatctcgtcc aacgtgatcc 60
tgctgctgctg 68

<210> SEQ ID NO 36
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 36
acttcatacc

<210> SEQ ID NO 37
<211> LENGTH: 96
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 
<222> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 37
agaacagtgt tcaagtgtgg agatgttga gatgggaaac actgtctctc acttcatacc
60
tgcgaatct ccacactga acaagttaga ccactt
96

<210> SEQ ID NO 38
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 
<222> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 38
agaacagtgt tcaagtgtgg agatgttga gatgggaaac actgtctctc acttcatacc
60
tgcgaatct cc
72

<210> SEQ ID NO 39
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 39
acaacgaa cagttgacc actt
24

<210> SEQ ID NO 40
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 40
atctccaa cgtgacaagt tagaccatt
30

<210> SEQ ID NO 41
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 41
casaacctca totatctctg ccatttaaat gcaatgtcac ggtaatggca gatagata
60
atgcaatgtc aaggttaa
77

<210> SEQ ID NO 42
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: 

<400> SEQUENCE: 42
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 42

caaactctta tcatatctcg ccattttaat gccttgcaca ggaatgcga ga

<210> SEQ ID NO 43
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 43
gtatgataat gcaatgtcac gccta

<210> SEQ ID NO 44
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 27
<223> OTHER INFORMATION: 3' Cy5
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 44
gtatgataat gcctttgacac gccta

<210> SEQ ID NO 45
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 45
tctgctattt cggctgcact gccttttaag tcagggact acgastgcac tggctcagga

<210> SEQ ID NO 46
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 46
tctgctattt cggctgcact gccttttaag tcagggact acgastgcac tg

<210> SEQ ID NO 47
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 47
tcaggtasg ttcaggtact acgaa

<210> SEQ ID NO 48
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 48
cattgcattc gtacctctg tgcgtttaag tatcagatcg ccgaagctac aggactcgaa

<210> SEQ ID NO 49
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 50
gactcgaag ttcagatcg ccgaa

<210> SEQ ID NO 51
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 51
cagctgatcgg cggatctga taacttttaat gaccaacca cctagtatc agatcgcgaa

<210> SEQ ID NO 52
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 52
cattgcattc ggctgatcga taacttttaat gaccaacca cctagtatc ag

<210> SEQ ID NO 53
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
SEQ ID NO: 54
LENGTH: 16
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic oligonucleotide

SEQUENCE: 54
atcgccgaat gaccaaac caccta 25

SEQ ID NO: 55
LENGTH: 16
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic oligonucleotide

SEQUENCE: 55
tgatactta ggtgtgtg tcacttttaa acactccta ctcaagtac ccaccaacta 60

SEQ ID NO: 56
LENGTH: 45
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: modified_base
LOCATION: 1
OTHER INFORMATION: 5' phosphorylation

SEQUENCE: 56
ttggtcattt taactccac tactcata cctcaagt gaccaacac caccta 45

SEQ ID NO: 57
LENGTH: 77
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic oligonucleotide

SEQUENCE: 57
aagctacag actacgaagt caatgtcag gtaagctaca ggtacagaa tttcgcgtg 60
acattgcaat attcatc 77

SEQ ID NO: 58
LENGTH: 25
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic oligonucleotide

SEQUENCE: 58
aagctacag actacgaagt caatg 25
<210> SEQ ID NO 59
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<222> LOCATION: 1
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 59
tcaggttaag ttagggact acgaatatta ccggttagtt gcaattacta tc 52

<210> SEQ ID NO 60
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 60
aagttacga tgcgggaag ttagggact acgaatatta ccggttagtt gcaattacta tttttcgt agt 60
ccggttagtt acgtgta 77

c<210> SEQ ID NO 61
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 61
aagttacga tgcgggaag ttagggact 25
tacag

c<210> SEQ ID NO 62
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<222> LOCATION: 1
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 62
gactacggaag ttagggact acgaatatta ccggttagtt gcaattacta tttttcgt agttaaccg 52
taaggtcaggtt tttttttaa gtaagttcgt gtaagttcgtgttaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtгапитовник 30
<210> SEQ ID NO 63
<211> LENGTH: 61
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 63
aagtacga acggtgtaag ttagggact acgaatatta ccggttagtt gcaattacta tttttcgt agttaaccg 60
taaggtcaggtt tttttttaa gtaagttcgt gtaagttcgtgttaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtgtaagttcgtгапитовник 61
<400> SEQUENCE: 64

ttggtcattc gggtgat

<400> SEQUENCE: 65

atgacaara ccaacctata ccaacctac tcaattttg tgtgt

<400> SEQUENCE: 66

atggtgcat gcagatc
gggtgat

<400> SEQUENCE: 67

aattggcagag atagataaga gtttg

<400> SEQUENCE: 68

aattggcagagt ggaatgtcaact acctcaactc gccatcatc cgtcatac
gggtgat

<400> SEQUENCE: 69

aattggcagag atagataaga gtttg

<400> SEQUENCE: 70

aattggcagag atagataaga gtttg
<400> SEQUENCE: 70

tgatgatgc gatgagg agtgcctc atcaactc atcaactt cactcgc a 59

<210> SEQ ID NO 71
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 71

tgatgatg cctc atcaactc atcaact t cactcgc a 49

<210> SEQ ID NO 72
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 5' phosphorylation

<400> SEQUENCE: 72
tcaactcgc a 10

<210> SEQ ID NO 73
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 73
gacatctc acgtcactc actactt 27

<210> SEQ ID NO 74
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: 3' 6-carboxyfluorescein

<400> SEQUENCE: 74
gacatctc acgtcactc actactt 27

<210> SEQ ID NO 75
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 75

ggtggctcta ggcgctgaa gtagtgattg agcgtagtga atgtcactc ttcaactcgc aatcactcag ctcactc 60

<210> SEQ ID NO 76
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 76

GTTAGTCTA GCAGCAGA GTAGATTG AGCGT  

75

<210> SEQ ID NO 77
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 77

GATAATGTC ACTACTCAA CTCGACTCA TCGTCTCA TC  

42

<210> SEQ ID NO 78
<211> LENGTH: 70
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 78

TCAAGAGC GCTAGACG ATAGAGCG TCTAGATA GCTAGCTT TGACTGCGC  

60

TAGACTACC  

70

<210> SEQ ID NO 79
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 79

GTCCAAGCT ATGACTACT ATGCACT  

27

<210> SEQ ID NO 80
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 80

GTCTATCCT GTOTGCGA GTAGTGATTG AGCGTGAGA ATGACTACT TCTAATCGC  

60

ATTCAATCG CTCAATC  

77

<210> SEQ ID NO 81
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 81

GTCTATCCT GTOTGCGA GTAGTGATTG AGCGT  

35

<210> SEQ ID NO 82
<211> LENGTH: 42
-continued

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5’ phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 92

gatgaatgct actacctcaa ctgcattca tcagctcaac tc

<210> SEQ ID NO 83
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 83

aaggtgctca tgaatggta ctctagct

<210> SEQ ID NO 84
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 84

gtcgatcatt cgtcgatgaa gtagtgattg agcggtgactagctagactac ttaaactgc

attcatcag ctcasct

<210> SEQ ID NO 85
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 85

gtcgattcct gtcgatgcaag ttagtattg agcctgactagct

<210> SEQ ID NO 86
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5’ phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 86

gatgaatgct actacctcaa ctgcattca tcagctcaac tc

<210> SEQ ID NO 87
<211> LENGTH: 129
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 87

cgtaatgc cagatagccc atgtgctcat cgatcagtcct aatctagtc 
tctggtggc cgcgtgagcag agctatcgt cctgatgac gaccagatac gagcgagctg

<210> SEQ ID NO 88
<211> LENGTH: 114
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 98

cgtaagcgc agagttagcctt cttgatgac gaccagatct cagcgtgagc gcagatgacta

tctggtggc cgcgtgagcag agctatcgt cctgatgac gaccagatac gagcgagctg

<210> SEQ ID NO 89
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' carboxytetramethylrhodamine (NHS Ester)
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 99
tacatcagca cgtagagctc gcotcctc

<210> SEQ ID NO 90
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' 6-carboxy-4', 5'dichloro-2', 7'-dimethoxyfluorescein (NHS Ester)
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 90
tacatcagca cgtagagctc gcotcctc

<210> SEQ ID NO 91
<211> LENGTH: 76
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 91

caggtgcgca gaggcttggag aaggtgctgagagctcag tcaacagcatcttcacgcga

tctcatacgc tcaactc
tctcatacgc tcaactc

<210> SEQ ID NO 92
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 92

caggtgcgca gaggcttggag aaggtgctgagagct

<210> SEQ ID NO 93
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 93

gatgaatgtc actacttca actcgcattc tcaagctca tc 42

<210> SEQ ID NO 94
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' carboxytetramethylrhodamine (NHS Ester)
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 94

actaacttag atctcgtagc tacgctg 27

<210> SEQ ID NO 95
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' 6-carboxy-4', 5'dichloro-2', 7'-dimethoxyfluorescein (NHS Ester)
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 95

actaacttag atctcgtagc tacgctg 27

<210> SEQ ID NO 96
<211> LENGTH: 77
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 96

atgagccaat ggcattgaga gtagtgattg agcgtagtga atgactacac ttcacatcgc 60
atccatcaac gcgctcactc 77

<210> SEQ ID NO 97
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide

<400> SEQUENCE: 97

atgagccaat ggcattgaga gtagtgattg agcg 35

<210> SEQ ID NO 98
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<210> SEQ ID NO 99
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 1
<223> OTHER INFORMATION: 5' phosphorylation
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 99

GATGAAATGTC ACTACTCAA CTGCGATCA TCAAGCTCAA TC

<210> SEQ ID NO 100
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 100

TTGCTCTGTA TCTAAACGA ACGAGCTCA GGCATTCA TCAAGCTCAA CACTACTT

<210> SEQ ID NO 101
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 101

TTGCTCTGTA TCTAAACGA ACGAGTCC

<210> SEQ ID NO 102
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: 29
<223> OTHER INFORMATION: 3' black hole quencher-1
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 102

AGGAGCTCA TCAAGCTCAA TCACTACTT

<210> SEQ ID NO 103
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide
<400> SEQUENCE: 103

CCTGATGAAT GTCTGGAGT CGTTGGTT
What is claimed is:

1. A system for performing a dynamic function comprising:
a first hairpin monomer comprising:
a first domain comprising a first toehold and a first propagation region, wherein the first toehold is exposed such that it is available to hybridize to a portion of a first nucleic acid sequence complementary to the first domain, and wherein the first toehold is located at an end of the monomer; and
a second domain comprising a second toehold and a second propagation region, wherein the second toehold is hybridized to a portion of the first propagation region;
an initiator molecule, wherein said first nucleic acid sequence complementary to the first domain comprises a portion of the initiator molecule; and
a second hairpin monomer comprising:
a first domain comprising a first toehold and a first propagation region, wherein the first toehold comprises a nucleic acid sequence complementary to the second toehold of the first monomer, wherein the first toehold is located at an end of the monomer and is exposed such that it is available to hybridize to the second toehold of the first hairpin monomer, and wherein the first propagation region is complementary to the second propagation region of the first hairpin monomer; and
a second domain comprising a second toehold and a second propagation region, the second toehold of the second hairpin monomer comprising a nucleic acid sequence complementary to the first toehold of the first hairpin monomer, wherein the second toehold of the second hairpin monomer is hybridized to a portion of the first propagation region of the second hairpin monomer, wherein the second propagation region of the second hairpin monomer comprises a sequence complementary to the first propagation region of the first hairpin monomer, and wherein the second domain of the second hairpin monomer comprises the sequence of the portion of the initiator that is complementary to the first domain of the first hairpin monomer so that upon
hybridization of the second domain of the second hairpin monomer to the first domain of the first hairpin monomer; any initiator hybridized to the first domain of the first hairpin monomer is displaced.

2. The system of claim 1, wherein said second toehold of the first hairpin monomer hybridizes to the first toehold of the second hairpin monomer if the first domain of the first hairpin monomer hybridizes to said initiator molecule.

3. The system of claim 1, wherein said second toehold of the first hairpin monomer initiates hybridization of said second propagation region of the first hairpin monomer to said first propagation region of said second hairpin monomer if the first domain of the first hairpin monomer hybridizes to said first nucleic acid sequence complementary to said first domain.

4. The system of claim 1, wherein said second propagation region of the first hairpin monomer comprises a portion of a single stranded hairpin loop.

5. The system of claim 1, wherein the first domain of the first hairpin monomer is an input domain and the second domain of the first hairpin monomer is an output domain.

6. The system of claim 1, wherein a portion of the first propagation region and the second toehold of the first hairpin monomer comprise a portion of a duplex stem.

7. The system of claim 1, wherein the first hairpin monomer further comprises a third domain comprising a third toehold and a third propagation region, wherein the third toehold is hybridized to a portion of the first propagation region of the first hairpin monomer.

8. The system of claim 1, wherein the first toehold of the first hairpin monomer is single stranded.

9. The system of claim 1, wherein said first domain and second domain of the first hairpin monomer are concatenated in the monomer.

* * * * *